



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le 2 Décembre 2016 par :
Sofiane Saadane

**Algorithmes stochastiques pour l'apprentissage, l'optimisation et
l'approximation du régime stationnaire**

JURY

FABIEN PANLOUP
SÉBASTIEN GADAT
PIERRE DEL MORAL
PEGGY CÉNAC
PIERRE-ANDRÉ ZITT
AURÉLIEN GARIVIER

Professeur des Universités
Professeur des Universités
Directeur de Recherche
Maitre de conférence
Maitre de conférence
Professeur des Universités

Directeur
Directeur
Rapporteur
Rapporteur
Examineur
Examineur

École doctorale et spécialité :

MITT : Domaine Mathématiques : Mathématiques appliquées

Unité de Recherche :

Institut de mathématiques de Toulouse (UMR 5219)

Directeur(s) de Thèse :

Sébastien Gadat et Fabien Panloup

Rapporteurs :

Pierre Del Moral et Peggy Cénac

Table des matières

1	Introduction Générale	13
1.1	Les algorithmes stochastiques	14
1.1.1	Contexte et premier résultat important	14
1.1.2	La méthode de l'ODE	16
1.2	Les algorithmes de bandit de Narendra et Shapiro	20
1.2.1	Les algorithmes de bandit stochastiques	21
1.2.2	Les algorithmes de Narendra et Shapiro	24
1.2.3	Le PMDM du bandit	30
1.3	L'algorithme de boule pesante stochastique	36
1.3.1	La descente de gradient	37
1.3.2	Descente de gradient à mémoire : Panorama des résultats connus	42
1.3.3	Présentation de nos résultats	46
1.4	Approximation du régime stationnaire du processus de McKean-Vlasov	57
1.4.1	Approximation du régime stationnaire de diffusions	58
1.4.2	Le processus de McKean-Vlasov	61
1.4.3	Les diffusions auto-agissantes	62
1.4.4	Approximation de la mesure invariante pour l'équation de McKean-Vlasov	65
1.4.5	Résultat	67
2	Regret bounds for Narendra and Shapiro bandit algorithms	71
2.1	Introduction	72
2.2	Definitions of the NS algorithms	75
2.2.1	Crude NSa and regret	75
2.2.2	Penalized and over-penalized two-armed NSa	76
2.2.3	Over-penalized multi-armed NSa	78
2.3	Main Results	78
2.3.1	Regret of the over-penalized two-armed bandit	78
2.3.2	Convergence of the multi-armed over-penalized bandit	80
2.3.3	Ergodicity of the limiting process	81
2.3.4	Total variation results	83
2.4	Proof of the regret bound (Theorems 2.3.1 and 2.3.2)	84
2.4.1	Notations	84
2.4.2	Evolution of $(Y_n)_{n \geq 1}$	85
2.4.3	Increase of exponent	87
2.4.4	Bound for $(\mathbb{E}(Z_n^{(3)}))_{n \geq n_0}$	90
2.4.5	Proof of Theorems 2.3.1 and 2.3.2 (a)	96

2.5	Almost sure and weak limit of the over-penalized bandit	98
2.5.1	A.s. convergence of the multi-armed bandit (Proposition 2.3.1)	98
2.5.2	Weak convergence of the normalized bandit (Proposition 2.3.2)	100
2.6	Ergodicity of the PDMP	102
2.6.1	Wasserstein results	102
2.6.2	Proof of total variation results	104
2.7	Technical result for the pseudo-regret upper bound	108
3	Stochastic Heavy Ball	111
3.1	Introduction	112
3.2	Stochastic Heavy Ball	113
3.2.1	Deterministic Heavy Ball	113
3.2.2	Stochastic HBF	114
3.2.3	Baseline assumptions	115
3.2.4	Main results	118
3.3	Almost sure convergence of the stochastic heavy ball	121
3.3.1	Preliminary result	121
3.3.2	Convergence to a local minimum	126
3.4	Convergence rates for strongly convex functions	136
3.4.1	Quadratic case	136
3.4.2	The non-quadratic case under exponential memory	142
3.5	Limit of the rescaled algorithm	145
3.5.1	Rescaling stochastic HBF	145
3.5.2	Tightness	146
3.5.3	Identification of the limit	148
3.5.4	Limit variance	152
3.6	Numerical experiments	155
3.7	Technical results	158
3.7.1	Standard tools of stochastic algorithms	158
3.7.2	Expectation of the supremum of the square of sub-Gaussian random variables	162
4	Computation of invariant distributions for stochastic McKean-Vlasov equations	169
4.1	Introduction	170
4.2	Preliminaries	174
4.3	Hypotheses	178
4.4	Main Result and Scheme of Proof	180
4.4.1	Tightness	181
4.4.2	Identification of the limit	182
4.5	Proofs of the results	184
4.5.1	Proof of Proposition 4.4.14	184
4.5.2	Proof of Theorem 4.4.2	190
4.6	Technical results	194
	Bibliographie	207

Remerciements

En trois ans de thèse, j'ai croisé beaucoup de personnes remarquables et beaucoup m'ont accompagné en m'encourageant ou en partageant de bons moments. Ces quelques lignes sont une occasion (bien trop rare) de les en remercier.

Une fois n'est pas coutume (deux fois non plus... En fait, on n'a jamais su au bout de combien fois c'était coutume...), je commencerai par remercier chaleureusement mes directeurs de thèse Fabien et Sébastien. Non seulement, vous m'avez offert un sujet de thèse passionnant mais j'ai aussi et surtout beaucoup appris à vos côtés. Sébastien, que je connais depuis la préparation à l'agrégation, et Fabien, qui a accepté d'encadrer et mon stage et ma thèse, votre bonne humeur, votre gentillesse et votre passion dans le travail rayonnent et sont un plaisir pour ceux qui vous côtoient au quotidien. Outre vos grandes qualités humaines, j'ai pu mesurer l'étendue de vos connaissances mathématiques au travers de nos nombreuses sessions de travail et apprendre le travail de chercheur dans des conditions idéales. Pour toutes ces raisons, je vous remercie du fond du coeur pour ces trois années.

Je tiens à remercier Peggy Cénac et Pierre Del Moral d'avoir accepté de rapporter ma thèse. Je suis aussi très heureux de compter Pierre-André Zitt et Aurélien Garivier parmi les membres de mon jury.

Une thèse c'est aussi un bureau avec des collègues. Je commencerai par remercier les membres permanents de l'IMT et plus particulièrement les membres de l'équipe ESP pour leur accueil et leur gentillesse.

Je garde le meilleur avec les doctorants de l'IMT. Tout d'abord, il me faut remercier les membres du bureau 203 et plus particulièrement Boubou, Yurko et Gigi. J'en ai vu des choses bizarres dans ma vie mais entrer dans ce bureau, c'est voyager en terre inconnue. J'ai vu : un homme téléphoner avec sa chaussure, un ukrainien faire de la lutte avec un mexicain, une fille sursauter à la moindre ouverture de porte, des bras de fer très virils, un homme rater sa bouche en buvant son café (Non je ne donnerai pas de nom.....c'est raf), un homme parler de *renié et moulottes*, un homme appeler la chambre des commerces pour prendre rendez-vous avec un médecin, une personne chasser la fameuse antenne du rectorat de l'université. J'ai surtout passé trois superbes années à vos côtés avec une immense collection de bons moments à jamais graver en mémoire. S'ils ne sont pas membres du bureau Stéphane et Kevin l'ont fréquenté à de nombreuses reprises pour des parties de coinches endiablées (où on a *compté les points*), des pauses ou des discussions à rallonges. Tous les 5, je vous remercie du fond coeur pour ces belles années. Je remercie aussi tout particulièrement Tatoche, Claire D., Myriam et Malika pour la gentillesse et bonne humeur dont elles ont fait preuve. Une pensée pour la nouvelle vague de doctorants qui a pris *le relais* : Valentin, Pierre, Ioana, Maylis, Daniel (avec qui j'ai le plaisir de partager mon bureau à l'INSA cette année) et Antoine (le poète du bureau 201). *Bon chance* à vous pour la suite.

Je tiens aussi à remercier Agnès et Isabelle pour les nombreux cafés et repas partagés toujours dans la bonne humeur.

Je tiens aussi à remercier mes amis proches. Merci à vous Alex, Cousin et Mathieu pour tous ces bons moments passés depuis notre entrée à la fac et pour votre amitié qui m'est très précieuse. Merci à Brendan pour sa bretonnerie et ses conseils musicaux toujours inspirés. Merci à Mélanie D. et Anaïs de supporter mes vanes depuis déjà trop longtemps.

Je terminerai par remercier mes parents et mes deux soeurs pour tout. Cette thèse vous est dédiée.

Avant-Propos

Dans cette thèse, nous étudions des thématiques autour des algorithmes stochastiques et c'est pour cette raison que nous débuterons ce manuscrit par des éléments généraux sur ces algorithmes en donnant des résultats historiques pour poser les bases de nos travaux. Ensuite, nous étudierons un algorithme de bandit issu des travaux de Narendra et Shapiro (voir [94]) dont l'objectif est de déterminer parmi un choix de plusieurs sources laquelle profite le plus à l'utilisateur en évitant toutefois de passer trop de temps à tester celles qui sont moins performantes. Notre but est dans un premier temps de comprendre les faiblesses structurelles de cet algorithme pour ensuite proposer une procédure optimale pour une quantité qui mesure les performances d'un algorithme de bandit, le regret. Dans nos résultats, nous proposerons un algorithme appelé NS sur-pénalisé qui permet d'obtenir une borne de regret optimale au sens minimax au travers d'une étude fine de l'algorithme stochastique sous-jacent à cette procédure. Un second travail sera de donner des vitesses de convergence pour le processus apparaissant dans l'étude de la convergence en loi de l'algorithme NS sur-pénalisé. La particularité de l'algorithme est qu'il ne converge pas en loi vers une diffusion comme la plupart des algorithmes stochastiques mais vers un processus à sauts non-diffusif ce qui rend l'étude de la convergence à l'équilibre plus technique. Nous emploierons une technique de couplage afin d'étudier cette convergence.

Le second travail de cette thèse s'inscrit dans le cadre de l'optimisation d'une fonction au moyen d'un algorithme stochastique. Nous étudierons une version stochastique de l'algorithme déterministe de boule pesante avec amortissement. La particularité de cet algorithme est d'être articulé autour d'une dynamique qui utilise une moyennisation sur tout le passé de sa trajectoire. La procédure fait appelle à une fonction dite de *mémoire* qui, selon les formes qu'elle prend, offre des comportements intéressants. Dans notre étude, nous verrons que deux types de mémoire sont pertinents : les mémoires exponentielles et polynomiales. Nous établirons pour commencer des résultats de convergence dans le cas général où la fonction à minimiser est non-convexe. Dans le cas de fonctions fortement convexes, nous obtenons des vitesses de convergence optimales en un sens que nous définirons. Enfin, l'étude se termine par un résultat de convergence en loi du processus après une bonne renormalisation.

La troisième partie s'articule autour des algorithmes de McKean-Vlasov qui furent introduit par Anatoly Vlasov et étudié, pour la première fois, par Henry McKean dans l'optique de la modélisation de la loi de distribution du plasma. Notre objectif est de proposer un algorithme stochastique capable d'approcher la mesure invariante du processus. Les méthodes pour approcher une mesure invariante sont connues dans le cas des diffusions et de certains autre processus mais ici la particularité du processus de McKean-Vlasov est de ne pas être une diffusion linéaire. En effet, le processus a de la mémoire comme les processus de boule pesante. De ce fait, il nous faudra développer une méthode alternative pour contourner ce problème. Nous aurons besoin d'introduire la notion de pseudo-trajectoires afin de proposer une procédure efficace.

Plan du manuscrit

Le chapitre 1 commence par des généralités sur les algorithmes stochastiques avec des rappels sur les résultats historiques classiques du domaine. Ensuite, nous proposons une vue d'ensemble des travaux en offrant au lecteur une grille de lecture de la thèse au travers d'une présentation des résultats principaux obtenus ainsi que des guides de lectures pour les preuves des résultats principaux.

Le chapitre 2 est constitué de l'article *Regret Bounds for Narendra and Shapiro Bandit Algorithms* écrit en collaboration avec Sébastien Gadat et Fabien Panloup.

Le chapitre 3 est constitué de l'article *Stochastic Heavy Ball* écrit en collaboration avec Sébastien Gadat et Fabien Panloup.

Le chapitre 4 est constitué d'un article en cours de rédaction à propos de l'approximation de la mesure invariante pour les équations de McKean-Vlasov.

Chapitre 1

Introduction Générale

Sommaire

1.1	Les algorithmes stochastiques	14
1.1.1	Contexte et premier résultat important	14
1.1.2	La méthode de l'ODE	16
1.2	Les algorithmes de bandit de Narendra et Shapiro	20
1.2.1	Les algorithmes de bandit stochastiques	21
1.2.2	Les algorithmes de Narendra et Shapiro	24
1.2.3	Le PMDM du bandit	30
1.3	L'algorithme de boule pesante stochastique	36
1.3.1	La descente de gradient	37
1.3.2	Descente de gradient à mémoire : Panorama des résultats connus	42
1.3.3	Présentation de nos résultats	46
1.4	Approximation du régime stationnaire du processus de McKean-Vlasov	57
1.4.1	Approximation du régime stationnaire de diffusions	58
1.4.2	Le processus de McKean-Vlasov	61
1.4.3	Les diffusions auto-agissantes	62
1.4.4	Approximation de la mesure invariante pour l'équation de McKean-Vlasov	65
1.4.5	Résultat	67

1.1 Les algorithmes stochastiques

Dans cette thèse, les trois thèmes qui seront abordés s'articulent autour de la notion d'algorithme stochastique. Avant de présenter nos travaux et résultats, nous commençons par définir ce qu'est un algorithme stochastique en toute généralité en proposant notamment quelques résultats historiques classiques.

1.1.1 Contexte et premier résultat important

Un algorithme stochastique est une procédure récursive dont le but est la recherche du zéro d'une fonction accessible seulement au travers d'évaluations bruitées et généralement sans biais. Cet objectif est à rapprocher de la méthode de Newton qui fut l'une des premières méthodes déterministes conçue pour trouver le minimum d'une fonction dont on ne connaît que peu d'informations. Imaginons pour commencer, que l'on cherche à minimiser une fonction h accessible seulement au travers d'une intégrale :

$$\forall x \in \mathbb{R}^d, \quad h(x) = \int_{\mathbb{R}^d} h(x, y) \nu(dy)$$

où ν est une mesure de probabilité sur \mathbb{R}^d . Si l'on applique un algorithme déterministe, le calcul de l'intégrale peut être trop coûteux si l'on a pas de formule explicite pour la fonction h . Ainsi on peut recourir à une formulation par algorithme stochastique en remarquant que si on dispose d'une suite i.i.d de variables aléatoires $(Z_n)_{n \geq 1}$ de loi ν alors l'algorithme s'écrit sous la forme suivante

$$X_{n+1} = X_n - \gamma_{n+1} h(X_n, Z_{n+1})$$

où $(\gamma_n)_{n \geq 1}$ est la suite de pas de l'algorithme. Généralement, c'est une suite décroissante qui tend vers 0 quand n tend vers l'infini et qui vérifie

$$\sum_{n \geq 1} \gamma_n = \infty.$$

L'intérêt de cette procédure se trouve dans la réécriture suivante. Posons

$$\Delta M_{n+1} = h(X_n) - h(X_n, Z_{n+1})$$

on obtient alors

$$X_{n+1} = X_n - \gamma_{n+1} h(X_n) + \gamma_{n+1} \Delta M_{n+1}$$

en remarquant que

$$\mathbb{E}(\Delta M_{n+1} | \mathcal{F}_n) = 0$$

c'est-à-dire que la suite $(\Delta M_n)_{n \geq 1}$ est une suite d'incrément de martingales. La particularité de cette écriture est de transformer notre équation de départ en un schéma d'Euler à pas décroissant $(\gamma_n)_{n \geq 1}$ et ainsi de donner une interprétation en lien avec la théorie des équations différentielles ordinaires.

Remarque Il est à noter que souvent h est le gradient d'une fonction *i.e.* $h = \nabla U$ et l'on retrouve ainsi l'algorithme dit de *gradient stochastique*. Dans la littérature, on parle souvent de *descente de gradient stochastique* mais le terme de descente n'est pas adapté car l'algorithme n'est plus vraiment en lien avec le schéma déterministe à cause de l'aléa.

Quelques unes des questions naturelles qui se posent sont alors les suivantes :

- la suite $(X_n)_{n \geq 1}$ converge-t-elle ? Si oui vers un minimum de h ?
- Quelle est la vitesse de convergence ?
- Peut-on envisager un théorème central limite ?

Historiquement, les premiers travaux sur les algorithmes stochastiques sont dus à Robbins et Monro dans les années 50 (voir [112]) et Kiefer-Wolfowitz (voir [69]) avec une approche utilisant une méthode de différences finies. Ces travaux ont donné naissance à une littérature très dense (voir par exemple [72]) car elles sont intéressantes non seulement d'un point de vue théorique mais aussi et surtout d'un point de vue applications pratiques. On peut notamment citer le travail de Bercu et Fraysse (voir [26]) pour le traitement des signaux des ECG, les travaux de Bubeck pour l'apprentissage séquentiel (voir [3]) ou encore en finance les travaux de Pagès et al. [79].

La première question qui se pose est celle de la convergence pour un algorithme stochastique. Nous allons ici donner deux méthodes, la première étant celle issue du lemme de Robbins-Siegmund qui se base sur une méthode de martingale et la seconde est basée sur une approche issue des systèmes dynamiques. Nous allons nous placer sur \mathbb{R} et rappeler quelques résultats de convergence pour le schéma suivant

$$X_{n+1} = X_n - \gamma_{n+1}f(X_n) + \gamma_{n+1}\Delta M_{n+1} \quad (1.1.1)$$

THEOREME 1.1.1. (Robbins-Siegmund) *On considère une filtration $(\mathcal{F}_n)_{n \geq 1}$ et quatre suites de variables aléatoires $(a_n)_{n \geq 1}$, $(b_n)_{n \geq 1}$, $(c_n)_{n \geq 1}$ et $(d_n)_{n \geq 1}$ qui sont $(\mathcal{F}_n)_{n \geq 1}$ -adaptées, **positives** et intégrables vérifiant :*

- (i) $(a_n)_{n \geq 1}$, $(b_n)_{n \geq 1}$, $(c_n)_{n \geq 1}$ et $(d_n)_{n \geq 1}$ sont des suites prévisibles.
- (ii) $\prod_{n \geq 1} (1 + a_n) < \infty$ et $\sum_{n \geq 1} \mathbb{E}(b_n) < \infty$.
- (iii) $\forall n \in \mathbb{N}$,

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n(1 + a_n) + b_n - c_n$$

Alors,

- (a) X_n converge p.s et $\sup_{n \geq 1} \mathbb{E}(X_n) < \infty$.
- (b) $\sum_{n \geq 1} \mathbb{E}(c_n) < \infty$ et $\sum_{n \geq 1} c_n < \infty$.

Remarques L'idée du résultat est de construire une surmartingale en utilisant (iii). Il est possible d'obtenir la convergence vers un minimum de f en construisant une fonction de Lyapunov. C'est-à-dire une fonction V sous-quadratique, positive et de classe C^2 qui tend vers l'infini en l'infini et alors d'appliquer le résultat de Robbins-Siegmund au contexte (1.1.1) pour obtenir le résultat de convergence vers un minimum. On notera que le résultat permet d'obtenir la convergence de $(X_n)_{n \geq 1}$ (sous de bonnes hypothèses) mais ne donne aucune idée de la vitesse de cette convergence. Sous des hypothèses plus fortes que celle du théorème de Robbins-Siegmund, on obtient la convergence presque sûre vers un minimum de f .

THEOREME 1.1.2. (Robbins-Monro) *Supposons que $(X_n)_{n \geq 1}$ satisfasse (1.1.1) et que*

- (i) f est continue.
- (ii) x^* est l'unique minimum de f .
- (iii) $\forall x \neq x^*$, $(x - x^*)f(x) \geq 0$.
- (iv) $(\Delta M_{n+1})_{n \geq 1}$ est une suite d'accroissements d'une martingale telle que pour tout entier n :

$$\mathbb{E}(|\Delta M_{n+1}|^2|\mathcal{F}_n) \leq C(1 + f(X_n)).$$

- (v) $\sum_{n \geq 1} \gamma_n = \infty$ et $\sum_{n \geq 1} \gamma_n^2 < \infty$.

Alors X_n converge p.s vers x^* .

1.1.2 La méthode de l'ODE

La méthode des systèmes dynamiques consiste à remarquer qu'un algorithme stochastique est un schéma d'Euler à pas décroissant. En effet, reprenons l'équation (1.1.1), divisons par γ_{n+1} et faisons tendre n vers l'infini. On obtient alors une approximation de l'équation différentielle suivante

$$\dot{x} = f(x) \quad (1.1.2)$$

Cette interprétation donne à penser que le comportement de la suite $(X_n)_{n \geq 1}$ est à rapprocher de celui l'ODE définie ci-dessus. C'est l'idée des travaux de Ljung (voir [86]) qui furent ensuite popularisés par le livre de Kushner et Clark en 1977 (voir [72]). On pourra citer le travail de Benaim (voir [20]) dans l'étude des systèmes dynamiques qui donna lieu à la théorie des pseudo-trajectoires (en collaboration avec Hirsch (voir [22])). Ceci offre ainsi une compréhension poussée du comportement asymptotique des processus stochastiques. Nous présentons ici le résultat de Kushner et Clark (voir [72]), fondateur dans la relation entre systèmes dynamiques et algorithmes stochastiques.

On commence par définir les notions d'ensemble invariant et de région d'attraction.

DEFINITION 1.1.1. Un ensemble Γ est dit invariant pour (1.1.2) si $x(0) \in \Gamma$ alors $x(t) \in \Gamma$ pour tout $t \geq 0$.

DEFINITION 1.1.2. On considère (1.1.2). Soit x^* un zéro de f . Un ensemble Γ est une région d'attraction pour x^* si

- Γ est dit invariant pour l'ODE.
- $x(0) \in \Gamma$ implique $\lim_{t \rightarrow \infty} x(t) = x^*$.
- Pour tout $\epsilon < 0$, il existe $\delta < 0$ tel que pour tout $x(0) \in \Gamma$ tel que $|x(0) - x^*| \leq \delta$ implique $|x(t) - x^*| \leq \epsilon$ pour tout $t \geq 0$. (x^* est alors dit asymptotiquement stable)

Il faut donc un moyen pour montrer qu'un ensemble Γ est une région d'attraction et la proposition suivante offre une façon de faire.

PROPOSITION 1.1.1. Soit Γ un ensemble invariant pour l'ODE (1.1.2) et x^* un zéro de f appartenant à l'adhérence de Γ . Supposons qu'il existe une fonction V de classe C^1 telle que

- $V(x) > V(x^*) = 0$ pour tout $x \neq x^*$.
- Γ est borné ou $\lim_{|x| \rightarrow \infty} V(x) = \infty$.
- Pour tout $x \in \bar{\Gamma} \setminus \{x^*\}$, on a $\langle \nabla V(x), f(x) \rangle > 0$

Alors Γ est une région d'attraction pour x^* .

Pour obtenir un résultat de convergence, l'idée est ensuite de construire un processus interpolé issue de l'algorithme (1.1.1).

DEFINITION 1.1.3. On appelle processus interpolé associé à (1.1.1) le processus défini comme suit

$$X^{(0)} = 0, \quad X^{(0)}(\Gamma_n) = X^{(0)}(\gamma_1 + \dots + \gamma_n) = X_n$$

avec

$$\Gamma_n = \sum_{k=1}^n \gamma_k, \quad \text{avec } n \geq 1.$$

Puis on interpole entre chaque instant de discrétisation

$$\forall t \in [\Gamma_n, \Gamma_{n+1}], \quad X^{(0)}(t) = X_n + (t - \Gamma_n)(-\Delta f(X_n) + \Delta M_{n+1})$$

On peut ensuite définir la fonction shift comme suit

$$\forall t \geq 0, \quad X^{(n)}(t) = X^{(0)}(t + \Gamma_n)$$

Cette définition est à voir comme une interpolation affine par morceaux aux points (Γ_n, X_n) . Introduisons une quantité importante appelée compteur d'itérations.

DEFINITION 1.1.4. Soit $T > 0$ et $n \in \mathbb{N}$, on appelle compteur d'itérations la quantité qui mesure le nombre d'itérations minimales nécessaires pour dépasser un temps T à partir d'un temps Γ_n :

$$N(n, T) = \inf\{k \geq n, \gamma_n + \dots + \gamma_{k+1} \geq T\}$$

On peut alors énoncer le résultat de Kushner et Clark.

THEOREME 1.1.3. On considère le schéma stochastique (1.1.1). On suppose que ∇f est continue et :

- (i) $(X_n)_{n \geq 1}$ est une suite bornée.
- (ii) Pour tout $T > 0$, $\lim_{n \rightarrow \infty} \sup_{j \leq N(n, T)} \left| \sum_{k=n}^j \gamma_k \Delta M_k \right| = 0$.

Alors,

- (a) La suite $(X_n)_{n \geq 1}$ est relativement compacte pour la topologie de la convergence uniforme sur les compacts de $C(\mathbb{R}^+, \mathbb{R}^d)$. De plus, toute valeur d'adhérence de cette suite est solution de l'équation différentielle.
- (b) Si Γ est une région d'attraction pour x^* et si $(X_n)_{n \geq 1}$ retourne infiniment dans Γ alors X_n converge vers x^* .

Nous venons de voir que pour obtenir la convergence presque sûre, il y avait plusieurs méthodes. Cependant, les résultats énoncés ne donnent pas d'idée de la vitesse de convergence. Ainsi, nous terminons par un résultat qui donne la vitesse de convergence de X_n vers sa cible x^* tiré de [25].

THEOREME 1.1.4. Supposons que les hypothèses du Théorème 1.1.2 sont satisfaites et que de plus

- (vi) f est deux fois continument différentiable.
- (vii) Il existe une matrice C définie positive telle que

$$\mathbb{E}(\Delta M_{n+1} \Delta M_{n+1}^T | \mathcal{F}_n) = C + o_p(1).$$

- (viii) Il existe $\delta > 0$ tel que

$$\sup_{n \geq 1} \mathbb{E}(|\Delta M_{n+1}|^{2+\delta} | \mathcal{F}_n) < \infty.$$

Alors, si on pose $\gamma_n = 1/n$ on a

$$\sqrt{n}(X_n - x^*) \rightarrow \mathcal{N}(0, \Sigma_f)$$

où

$$\Sigma_f = \int_0^\infty \exp(-(Df(x^*) - \frac{Id}{2})s)^T C \exp(-(Df(x^*) - \frac{Id}{2})s) ds$$

Remarques En particulier, ce résultat donne une vitesse de convergence pour un pas de la forme $\gamma_n \simeq 1/n$. On notera que les travaux de Ruppert et Polyak (voir [113]) ont montré qu'une moyennisation permet d'obtenir une vitesse de convergence pour un pas différent de $1/n$. L'idée est à chaque étape de considérer la moyenne des X_n définis par (1.1.1) et donc de considérer la suite

$$\bar{X}_0 = X_0, \quad \bar{X}_n = \frac{1}{n} \sum_{k=0}^n X_k$$

Précisons que dans cette définition le pas n'est pas forcément $\gamma_n = 1/n$. Ainsi on obtient un nouvel algorithme

$$\bar{X}_{n+1} = \bar{X}_n + \frac{1}{n+1} (X_{n+1} - \bar{X}_n)$$

Polyak et Ruppert (voir [113]) ont montré que l'on obtient ainsi un estimateur asymptotiquement efficace de x^* pour un pas $\gamma_n = n^{-\beta}$ avec $\beta \in (1/2, 1)$.

Comme nous le verrons dans cette thèse, pour obtenir une vitesse, il faut obtenir un contrôle plus fin de l'algorithme. Nous allons voir que l'on peut obtenir une vitesse si l'on parvient au préalable à obtenir une inégalité de la forme,

$$\forall n \geq n_0, \quad \mathbb{E}(f(X_{n+1}) | \mathcal{F}_n) \leq f(X_n)(1 - \gamma_n) + C\gamma_n^2$$

pour le schéma (1.1.1) où $C \geq 0$. On obtient alors, par une récurrence immédiate,

$$\mathbb{E}(f(X_n)) \leq \mathbb{E}(f(X_0)) \prod_{k=1}^n (1 - \gamma_k) + C \sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - \gamma_l)$$

Pour contrôler ces termes, nous verrons que les principaux outils sont des comparaisons entre séries et intégrales en utilisant les pas suivant

$$\gamma_n = \gamma n^{-\beta}$$

avec $\beta \in [0, 1]$ et $\gamma > 0$. On obtient alors une borne de la forme suivante :

$$\mathbb{E}(f(X_n)) \leq Cn^{-\theta}$$

où θ est un réel compris entre 0 et 1 dépendant de β et γ .

Remarque Nous verrons qu'obtenir de telles bornes est difficile et demande des propriétés de régularité supplémentaires sur la fonction f . Terminons en citant quelques ouvrages parmi les plus complets sur les algorithmes stochastiques : [72] qui offre un panorama sur la méthode de l'EDO, [73] est une référence plus récente et donc plus complète sur le domaine, [44] propose une partie sur le contournement des pièges pour les procédures stochastiques (utile notamment pour notre travail sur les algorithmes de boule pesante) et le séminaire de probabilité de Benaïm (voir [20]) qui offre

une compréhension poussée du comportement limite des algorithmes stochastiques en traitant de la notion de pseudo-trajectoire asymptotique.

1.2 Les algorithmes de bandit de Narendra et Shapiro

Le terme *bandit* tire son origine des machines à sous que l'on trouve dans les casinos. Le fonctionnement d'une machine se résume à actionner un bras qui engendre deux types de résultats : un gain ou (comme très souvent) une perte. Plus généralement, on parle d'algorithme de bandit en évoquant une procédure qui a pour but de déterminer parmi un panel de plusieurs sources celle qui sera la plus profitable au joueur. Une telle procédure sera certainement efficace si elle ne passe pas trop de temps à utiliser les moins bonnes.

Les algorithmes de bandits sont ainsi des problèmes d'allocations séquentielles de ressources dont les applications sont nombreuses. La plus ancienne est sans doute celle qui donne lieu aux travaux de Thompson [1] pour les tests cliniques. Imaginons qu'un patient vienne consulter un médecin et que ce dernier, hésitant, ait à sa disposition cinq traitements. Construire une procédure afin de déterminer le traitement le plus efficace rentre dans le cadre des problèmes de bandits avec une contrainte humaine naturelle car il faut éviter de trop tester de médicaments sur le patient. Parmi les autres applications, on peut penser aux publicités sur internet dont le fonctionnement repose sur le nombre de clics effectués sur une fenêtre apparaissant sur l'écran de l'utilisateur. On trouve aussi des liens avec la thématique du routage de données.

D'un point de vue purement mathématique, les algorithmes de bandits peuvent être formalisés d'au moins trois façons : les bandits stochastiques, les bandits adversaires et les bandits Markoviens. La principale différence entre ces familles se situe dans la manière dont une action judicieuse est récompensée. Dans notre travail, nous ne nous intéresserons qu'à la seule famille des bandits stochastiques et nous renvoyons aux travaux de Bubeck et Cesa-Bianchi [34] proposant un panorama complet sur les problèmes de bandit et les résultats majeurs du domaine. Dans la vaste littérature du monde des bandits stochastiques, il existe de nombreux algorithmes dont les stratégies varient en fonction de l'objectif en vue. Pour comprendre ces différences, il faut définir un moyen de mesurer l'efficacité d'un algorithme. Le regret est généralement une quantité d'intérêt, il mesure l'écart entre la récompense cumulée obtenue si on avait à chaque étape opté pour le meilleur bras et la récompense cumulée obtenue par la stratégie mise en place par le joueur. Notre étude porte donc sur une classe d'algorithmes de bandits pour lesquels nous proposons une étude fine du regret au moyen d'une étude détaillée d'un algorithme stochastique.

Un autre aspect important qui sera abordé dans cette partie est l'étude de **P**rocessus de **M**arkov **D**éterministes par **M**orceaux, nom qui sera maintenant raccourci via l'acronyme PMDM. En effet, habituellement un algorithme stochastique renormalisé converge vers un processus de diffusion (voir [73] pour plus de détails). Une fois renormalisé judicieusement les algorithmes de Narendra et Shapiro convergent en loi vers un processus présentant une structure particulière. Ici, cette limite se caractérise par une trajectoire déterministe la plupart du temps et par des sauts à des temps aléatoires. Ce type de processus a été introduit par Davis dans les années 80 et donne lieu depuis une dizaine d'années à une littérature dense et variée. Un PMDM se caractérise par trois éléments : le flot déterministe qui détermine la trajectoire entre les sauts, le taux de saut qui détermine l'intensité des sauts et le noyau de transition qui détermine la position après un saut. Les études de tels processus interviennent dans plusieurs domaines : biologie, gestion du flux de réseaux de communications ou encore la fiabilité d'automates pour n'en citer que quelques-unes (voir l'article de survol de [9] et les références qu'il contient). Notre étude se focalisera sur la vitesse de convergence vers l'état d'équilibre. Nous proposerons des méthodes de couplage qui jouent sur

la structure du processus et permettent notamment d’obtenir des résultats pour des distances spécifiques (principalement les distances de Wasserstein et en Variation Totale).

Contenu de la sous-partie

Dans cette partie, notre objectif est de présenter les résultats de l’article *Regret bounds for Narendra and Shapiro bandits algorithm*. Dans cette optique, nous commençons par présenter la notion d’algorithme de bandit et celle de regret comme moyen de quantifier les performances de telles procédures. La particularité de notre travail est d’offrir pour la première fois une étude du regret par le biais de celle d’un algorithme stochastique. Nous verrons que nous parvenons à obtenir des bornes optimales en un sens que nous définirons plus loin. Inhérente à l’étude des algorithmes stochastiques, la convergence en loi de l’algorithme sera un second objectif de l’article et nous verrons que la limite est un PMDM pour lequel nous proposerons une série de résultats accés sur sa vitesse de convergence à l’équilibre.

1.2.1 Les algorithmes de bandit stochastiques

Nous définissons en toute généralité la procédure associée à un algorithme de bandit.

Introduit pour la première fois par les travaux de Robbins [111] en 1952 pour les bandits stochastiques, la formalisation du problème de bandit comme un jeu contre un adversaire s’est développée sous l’impulsion des travaux de Auer et al. (voir [5]) et depuis une dizaine d’années de nombreux auteurs se sont attaqués à ces problèmes avec un intérêt à la fois théorique mais aussi tourné vers les applications modernes notamment dans le domaine du machine learning. Parmi les applications nous trouvons le placement de bandeaux publicitaires sur les pages internet par exemple. Ce modèle fonctionne sur le principe suivant : une série de publicités est proposée à l’utilisateur et chaque fois qu’il clique sur l’une de ces annonces un revenu est perçu et on lui en propose de nouvelles en fonction des clics effectués. L’objectif étant de maximiser le nombre de clics. Ici le nombre de bras est l’ensemble des publicités disponibles ainsi que leur emplacement sur la page et le temps correspond à un nouvel utilisateur se connectant au serveur (voir le travail de Pandey et al. [103] et Liu and Zhao [85] notamment). Un autre exemple est celui du transfert de données d’un utilisateur vers un autre. On considère un réseau où le but est d’envoyer des données d’un point du graphe vers un autre avec une contrainte temporelle. Le but est donc de trouver un chemin le moins gourmand en temps. Ici les bras sont les chemins offerts par le graphe et le temps est celui du routage entre deux noeuds successifs (voir Awerbuch and Kleinberg [8], McMahan et Blum [90], Gyorgy et al. [60]). On peut aussi citer les problèmes d’optimisation convexe en lien avec l’algorithme de descente miroir qui se modélise comme un problème de bandit (voir [34]).

Le fonctionnement d’un algorithme de bandit stochastique est relativement simple à comprendre. Nous parlerons d’algorithme de bandit à d bras lorsque un joueur aura le choix parmi d sources. A chacune de ces sources est associée une récompense notée A_i^n où $i \in \{1, \dots, d\}$ fait référence au bras choisi et n au nombre de fois où l’on a joué. Le terme *stochastique* signifie ici que la suite de récompense $(A_i^n)_{n \geq 1}$ (pour tout $i \in 1, \dots, d$) est une suite i.i.d de variables aléatoires i.e la loi de la récompense obtenue est indépendante du temps. Dans le cas qui nous intéresse les récompenses seront extrêmement simples et mimeront des variables aléatoires de Bernoulli de paramètre p_i , $i \in 1, \dots, d$. Sans perte de généralités, nous supposons qu’un seul des bras est optimal, c’est-à-dire qu’il existe un unique bras tel que $p_i > p_j$, $\forall i \neq j$. L’algorithme suivant décrit de manière générale le fonctionnement d’un algorithme de bandit

Algorithm 1 Procédure de bandit stochastique

- Paramètres connus : le nombre de bras d et (souvent) le nombre de fois où l'on joue N .
- Paramètres inconnus : les paramètres des lois de Bernoulli et (rarement) le nombre de fois où l'on joue N .

Principe du jeu :

- 1 Choisir un bras $i \in 1, \dots, d$.
 - 2 En fonction du choix effectué : le joueur reçoit une récompense et recommence (1).
-

Les différents algorithmes connus se distinguent par le choix des récompenses et surtout la manière dont on choisit le bras. On peut citer en exemple la famille des algorithmes UCB (Upper Confidence Bound) introduite par Auer et al. [6] dont le but est de construire à chaque étape une borne supérieure de confiance pour la moyenne de chaque récompense. Pour ce faire, à chaque bras est associé une quantité appelée index et on tire à chaque étape le bras ayant le plus grand index. De nombreux raffinements ont été introduits avec pour objectif l'obtention d'algorithmes les plus efficaces possibles. Nous donnons ci-dessous un algorithme pour illustrer cela. La procédure MOSS (Stratégie Minimax Optimale dans le cas Stochastique) est due à Audibert et Bubeck [3]. Son principe est le suivant :

Algorithm 2 MOSS

On définit au préalable N comme étant le nombre de fois où l'on joue. On note \hat{A}_i^n la moyenne empirique des récompenses du bras i après n étapes. Principe de jeu :

- Pour chaque bras i définir un index

$$B_{i,n} = \hat{A}_i^n + \sqrt{\frac{\max(\log(\frac{n}{dN}), 0)}{n}}$$

- On choisit ensuite le bras ayant le plus grand index.
-

Cet algorithme présente une force incontestable d'un point vue performance mais demande au préalable de connaître le nombre de fois où une action doit être effectuée ce qui n'est pas le cas des algorithmes que nous étudions. On parle ainsi de stratégie *anytime* pour une stratégie où le nombre de fois où l'on joue n'a pas besoin d'être prédéfini à l'avance.

Évidemment, il faut définir un moyen de jauger l'efficacité d'un algorithme de bandit. Avant de présenter les algorithmes qui vont nous intéresser, nous allons définir le regret qui permet d'évaluer les performances d'un algorithme de bandit. Comme nous le disions plus haut, le regret compare la récompense cumulée obtenue en utilisant une stratégie donnée et la récompense induite par le choix à chaque étape du bras optimal. On notera I_k le bras choisi au temps k par le joueur. Nous supposons dans la suite que le bras 1 est le bras optimal sans aucune perte de généralité. Le regret est la variable aléatoire définie comme suit :

$$R_n := \underbrace{\max_{1 \leq j \leq d} \sum_{k=1}^n [A_k^j]}_{\text{récompense maximale théorique}} - \underbrace{\sum_{k=1}^n A_k^{I_k}}_{\text{récompense obtenue}} .$$

Une bonne stratégie est donc une stratégie qui minimise cette quantité : plus le regret est petit et plus nous obtenons une récompense cumulée optimale. Les premiers résultats sur le regret sont

dus à Lai et Robbins (voir [75]) avec le résultat suivant.

THEOREME 1.2.5. *Soit $T_i(n)$ le nombre de fois où l'on a joué le bras i après n étapes, μ_i le paramètre de la i -ème loi de Bernoulli associée au bras i et $\Delta_i = \mu^* - \mu_i$ avec μ^* le paramètre du bras le plus fort. Pour toute stratégie vérifiant*

$$\mathbb{E}(T_i(n)) = o(n^a),$$

et pour tout algorithme de bandits avec pour récompenses des lois de Bernoulli, pour tout i tel que $\Delta_i > 0$ et tout $a > 0$, on a alors

$$\liminf_{n \rightarrow \infty} \frac{R_n}{\log(n)} \geq \sum_{i: \Delta_i > 0} \frac{\Delta_i}{kl(\mu_i, \mu^*)}$$

où kl désigne la divergence de Kullback-Leibler.

Dans le cas de récompense de type Bernoulli, Bubeck et Cesa-Bianchi ont donné une preuve simplifiée (voir [34]). Approcher cette quantité a motivé de nombreux travaux et beaucoup sont en lien avec les algorithmes UCB. Par exemple, l'algorithme UCB-V introduit par Audibert et al. (voir [34]) prend en compte la variance des récompenses ou encore KL-UCB introduit indépendamment par Garivier et Cappé (voir [37]) et Maillard et al. (voir [87]) atteint la borne de Lai et Robbins qui est de l'ordre de $\log(n)$. Notre objectif est ici différent car nous souhaitons obtenir une borne uniforme pour le regret. Nous parlerons de stratégie *minimax* pour désigner celle qui minimise le regret uniformément sur les probabilités associées à chaque bras et sur toutes les stratégies admissibles. Plus précisément, une stratégie qui conduit à un regret de l'ordre de $\inf \sup \mathbb{E} R_n$ (où le supremum est pris sur le vecteur $p = (p_1, \dots, p_d) \in (0, 1)^d$ et l'infimum sur toutes les stratégies possibles) seront minimax. Cette quantité est difficile à manipuler car il faut contrôler l'espérance du supremum d'un processus. Pour contourner ce problème, nous étudierons la quantité appelé pseudo-regret :

$$\bar{R}_n := \max_{1 \leq j \leq d} \mathbb{E} \sum_{k=1}^n \left[A_k^j - A_k^{I_k} \right].$$

La proposition qui suit justifie que l'étude du pseudo-regret suffit à déterminer l'ordre de grandeur du véritable regret. Le résultat suivant est extrait de [34].

PROPOSITION 1.2.2. *Considérons un algorithme de bandit stochastique dont les récompenses sont des variables aléatoires suivant une loi de Bernoulli de paramètre $p = (p_1, \dots, p_d) \in (0, 1)^d$. Nous avons alors uniformément en p :*

$$0 \leq \mathbb{E} R_n - \bar{R}_n \leq \sqrt{\frac{n \log d}{2}}.$$

De plus, si on considère l'infimum sur toutes les stratégies possibles et le supremum sur toutes les variables aléatoires de Bernoulli alors

$$\inf \sup \mathbb{E}[R_n] \geq \frac{1}{20} \sqrt{nd}.$$

On notera dans ce résultat que le terme $\sqrt{\frac{n \log d}{2}} = \sqrt{nd \frac{\log(d)}{2d}}$ est du même ordre en n que \sqrt{nd} avec une constante dépendante de d supplémentaire mais négligeable dans l'asymptotique en n .

Pour les algorithmes UCB, il est facile de montrer que l'algorithme $a-UCB$ a un regret borné par $\sqrt{an \log(n)}$ (voir [34]). Audibert et Bubeck ont proposé une modification permettant de supprimer le facteur logarithmique de cette borne (voir [3]). Citons aussi l'algorithme MOSS pour lequel il est prouvé que le regret est borné par $5,7\sqrt{nd}$. On notera aussi les travaux de Auer et Ortner (voir [7]) (voir aussi Perchet et Rigolet [107]) qui ont créé une stratégie UCB améliorée permettant d'obtenir un regret majoré par $\sum_{i:\Delta_i>0} \frac{\log(n\Delta_i^2)}{\Delta_i}$ qui, dans certaines configurations, fait mieux que

MOSS sans être toutefois minimax. Signalons aussi le travail de Garivier, Ménard et Stoltz (voir [54]) qui fournit des preuves simplifiées pour l'obtention de borne inférieure pour le regret en ne faisant appel qu'à certaines propriétés de la divergence de Kullback-Leibler. En résumé, cette brève revue bibliographique nous invite à la recherche d'une borne de l'ordre de \sqrt{nd} .

1.2.2 Les algorithmes de Narendra et Shapiro

L'objectif de cette sous-section est de présenter nos résultats. Nous commencerons par présenter les algorithmes d'intérêt avec une présentation des diverses modifications de l'algorithme de Narendra-Shapiro [94] qui aboutiront à une borne de regret minimax. Nous énoncerons les résultats sur le regret avant de passer dans une seconde partie sur les résultats propres aux PMDM.

Définitions et premiers résultats

Les algorithmes de Narendra et Shapiro tirent leurs origines dans une procédure établie dans les années 60 (voir [94]) en se posant comme un modèle intéressant d'algorithme stochastique. La procédure proposée par cet algorithme est simple : à chaque étape un bras est choisi et on procède à son évaluation. Si cette dernière est probante alors le bras est récompensé et sinon on ne fait rien. Afin d'offrir un cadre plus simple, nous supposerons que seulement 2 bras sont en compétition avec pour récompenses des lois de Bernoulli de paramètres respectifs p_1 et p_2 . De plus nous supposerons que p_1 est le bras optimal i.e $p_1 > p_2$ et noterons dorénavant $\pi := p_1 - p_2$.

Notons X_n la probabilité de choisir le bras 1 à l'étape n et $(\gamma_n)_{n \geq 0}$ une suite de pas décroissante qui tend vers 0 quand n tend l'infini. L'algorithme introduit par Narendra et Shapiro en 1969 sera appelé NS non pénalisé et défini comme suit :

$$X_{n+1} = X_n + \begin{cases} \gamma_{n+1}(1 - X_n) & \text{si le bras 1 est sélectionné et gagne.} \\ -\gamma_{n+1}X_n & \text{si le bras 2 est sélectionné et gagne.} \\ 0 & \text{sinon} \end{cases} \quad (1.2.3)$$

Notons que la construction est symétrique, i.e. $1 - X_n$ (qui correspond à la probabilité de choisir le bras 2) a une expression symétrique à celle écrite ci-dessus et suit donc une évolution duale de celle de X_n .

Cette algorithme possède un problème majeur : il n'est pas *infaillible*. Plus précisément, l'algorithme ne converge pas vers le bras optimal et échoue donc dans sa tâche initiale. Cette faiblesse s'explique par le fait que la procédure n'apprend pas de ses erreurs (voir dernier cas de (1.2.3)). Concrètement, lorsque qu'un bras perd, aucun enseignement n'est tiré par le joueur. Or, un algorithme de bandit efficace se doit de proposer un équilibre entre l'exploration des possibilités offertes (c'est-à-dire, jouer de temps en temps les bras sous-optimaux) et l'exploitation d'une action (il faut tenir compte des succès mais aussi des échecs). En particulier, la faillibilité de l'algorithme

(i.e. $\mathbb{P}(X_n \xrightarrow{n \rightarrow +\infty} 1) < 1$) génère un regret linéaire en le nombre de fois où l'on joue et est donc loin de la borne optimale comme le montre le calcul heuristique ci-dessous.

$$\begin{aligned}\bar{R}_n &= p_1 n - \sum_{k=1}^n (p_1 \mathbb{E}[X_k] + p_2 \mathbb{E}[1 - X_k]) = \pi \mathbb{E} \left[\sum_{k=1}^n (1 - X_k) \right] \\ &\gtrsim \pi \mathbb{P}(X_\infty = 0) \times n.\end{aligned}\tag{1.2.4}$$

Dans une optique d'amélioration des performances, Lambertson et Pagès (voir [79]) ont effectué un premier pas vers l'obtention d'un algorithme pertinent d'un point de vue regret. En effet, comme nous l'avons dit plus haut, l'algorithme de Narendra et Shapiro ne tient pas compte des échecs. En introduisant une pénalité, on parvient à obtenir un algorithme infaillible. Nous définissons l'algorithme suivant appelé NS pénalisé :

$$X_{n+1} = X_n + \begin{cases} \gamma_{n+1}(1 - X_n) & \text{si le bras 1 est sélectionné et gagne.} \\ -\gamma_{n+1}X_n & \text{si le bras 2 est sélectionné et gagne.} \\ -\rho_n \gamma_{n+1}X_n & \text{si le bras 1 est sélectionné et perd.} \\ \rho_{n+1} \gamma_{n+1}(1 - X_n) & \text{si le bras 2 est sélectionné et perd.} \end{cases}\tag{1.2.5}$$

Il est important ici de noter la nature des suites de pas $(\rho_n)_{n \geq 1}$ et $(\gamma_n)_{n \geq 1}$. En effet, suivant les résultats de [79], on peut choisir

$$\gamma_n = \frac{\gamma_1}{n^\alpha}, \quad \rho_n = \frac{\rho_1}{n^\beta}\tag{1.2.6}$$

avec $\alpha, \beta > 0$. L'apport majeur entre NS non pénalisé et l'algorithme pénalisé réside dans la capacité de l'algorithme à toujours désigner le bras optimal et ceci, sous des hypothèses minimales qui sont adaptatives : elles ne dépendent pas de quantité inconnues dans le problème de bandit. Le résultat suivant est tiré de [79].

PROPOSITION 1.2.3. *Soient $0 \leq p_2 < p_1 \leq 1$, $\gamma_n = \frac{\gamma_1}{n^\alpha}$, $\rho_n = \frac{\rho_1}{n^\beta}$ avec $\alpha, \beta > 0$ et $(\gamma_1, \rho_1) \in (0, 1)^2$. Notons $(X_n)_{n \geq 1}$ est la suite définie par (1.2.5).*

- (i) *Si $0 < \beta \leq \alpha$ et $\alpha + \beta \leq 1$, NS pénalisé est infaillible.*
- (ii) *De plus, si $0 < \beta \leq \alpha$ et $\alpha + \beta \leq 1$, alors $\frac{1-X_n}{\rho_n} \rightarrow \frac{1-p_1}{p_1-p_2}$.*

En dépit de sa capacité à toujours désigner le bras optimal, cette algorithme présente malgré tout une faiblesse structurelle de taille. En effet, son regret reste linéaire pour certaines configurations. Lorsque les bras ont des paramètres tous les deux proches de 1, l'algorithme a un regret de l'ordre de n où n est le nombre d'itérations de l'algorithme. Pour comprendre le problème, il faut étudier l'algorithme stochastique issu de la procédure (1.2.5). On considère :

$$X_{n+1} = X_n - \pi \gamma_{n+1} h(X_n) + \gamma_{n+1} \rho_{n+1} \kappa(X_n) + \gamma_{n+1} \rho_{n+1} \Delta M_{n+1}\tag{1.2.7}$$

où

$$\begin{aligned}h(x) &= x(1-x) \\ \kappa(x) &= -(1-p_1)x^2 + (1-p_2)(1-x)^2\end{aligned}$$

L'idée de notre preuve repose sur la définition du regret et une calibration optimale des pas de

l'algorithme.

Idée centrale de la preuve

Commençons par effectuer une renormalisation de l'algorithme. Nous rappelons que le bras 1 est supposé être l'unique bras optimale. Ainsi X_n tend vers 1 p.s en vertu de la Proposition 1.2.3. On considère donc la variable aléatoire $(Y_n)_{n \geq 1}$ définie comme suit :

$$Y_n = \frac{1 - X_n}{\rho_n}$$

Reprenons la définition du regret et utilisons la nouvelle variable introduite ci-dessus.

$$\bar{R}_n = \pi \mathbb{E} \left(\sum_{k=1}^n (1 - X_k) \right) = \pi \sum_{k=1}^n \rho_k \mathbb{E}(Y_k) \quad (1.2.8)$$

De cette formule, on peut tirer deux enseignements. Tout d'abord, si on prouve que $\sup_{n \geq 1} \mathbb{E}(Y_n) < \infty$ alors on aura

$$\bar{R}_n \leq C \sum_{k=1}^n \rho_k.$$

Ce qui nous amène au second point à savoir le calibrage des suites de pas de l'algorithme. En effet, nous savons (voir Proposition 1.2.3) que les pas doivent vérifier l'équation

$$\sum_{n \geq 1} \gamma_n \rho_n = \infty$$

pour être infaisible.

C'est-à-dire, en reprenant (1.2.6), que l'on doit avoir $\alpha + \beta \leq 1$. Rappelons que l'on cherche à obtenir un regret le plus petit possible. On veut donc minimiser $\sum_{k=1}^n \rho_k$ c'est-à-dire que l'on veut β le plus grand possible. On choisit donc $\alpha = \beta = 1/2$. On a alors

$$\sum_{k=1}^n \rho_k = \sum_{k=1}^n \frac{1}{\sqrt{k}} \leq C \int_1^n \frac{dx}{\sqrt{x}} = \tilde{C} \sqrt{n}.$$

On observe ainsi que cette calibration des pas conduit à l'ordre optimal voulu.

Ainsi, notre objectif est maintenant de montrer que

$$\sup_{n \geq 1} \mathbb{E}(Y_n) < \infty$$

en considérant les pas

$$\gamma_n \simeq \frac{1}{\sqrt{n}}, \quad \rho_n \simeq \frac{1}{\sqrt{n}}.$$

Les insuffisances de NS pénalisé et non pénalisé

La faiblesse structurelles des algorithmes NS pénalisé et non pénalisé s'explique par un rappel faible qui ne suffit pas à ramener l'algorithme vers le bras optimale. Afin de rendre cette intuition plus précise, définissons l'algorithme renormalisé.

$$Y_{n+1} = Y_n(1 + \gamma_{n+1}(\epsilon_n - \pi X_n)) - \gamma_{n+1}\kappa(X_n) + \gamma_{n+1}\Delta M_{n+1} \quad (1.2.9)$$

où $\epsilon_n = \frac{1}{\rho_{n+1} - \frac{1}{\rho_n}}$. Nous utiliserons la formulation suivante dorénavant

$$Y_{n+1} = Y_n + \gamma_n \varphi_n(Y_n) + \gamma_{n+1} \Delta M_{n+1} \quad (1.2.10)$$

avec

$$\varphi_n(y) = \underbrace{y(\epsilon_n - \pi(1 - \gamma_n y))}_{\varphi_n^1(y)} + \underbrace{\left(-\frac{\gamma_{n+1}}{\gamma_n} \kappa(1 - \gamma_n y)\right)}_{\varphi_n^2(y)}$$

La quantité ainsi définie joue un rôle prépondérant puisque son signe décide du comportement de $(Y_n)_{n \geq 1}$. Pour obtenir le résultat souhaité, il faut s'assurer que $\varphi_n(\cdot)$ reste négative ce qui assure une propriété de rappel. On notera que plus y est grand plus $\varphi_n(y)$ est grand. Plus simplement, ce terme assure la non-explosion de l'algorithme. Rappelons que X_n vit dans $[0,1]$ et que par conséquent Y_n vit dans $[0, \rho_n^{-1}]$. Pour l'algorithme NS non pénalisé on a alors

$$\varphi_n^1(y) < 0 \Leftrightarrow x \geq \frac{\epsilon_n}{\pi}$$

ce qui signifie que la dérive se comporte mal au voisinage de 0. En particulier, NS non pénalisé ne peut fournir une borne de regret intéressant et en se tournant vers l'algorithme NS pénalisé on se rend compte que c'est le terme $\varphi_n^2(\cdot)$ qui va aider la dynamique pour des valeurs de X_n proche de 0. En effet,

$$\varphi_n^2(\rho_n^{-1}) = \frac{1}{2\gamma_1^2} - \frac{\gamma_1}{\rho_1}(1 - p_2) < 0 \Leftrightarrow p_2 < 1 - \frac{\rho_1}{2\gamma_1^3}$$

La figure 1.2.2 illustre ces considérations. La proposition suivante résume tout cela de manière plus précise.

THEOREME 1.2.6. *Soit $(X_n)_{n \geq 0}$ la suite définie par l'algorithme du bandit pénalisé (1.2.5) avec $(\gamma_n, \rho_n)_{n \geq 1}$ et $(\gamma_1, \rho_1) \in (0,1)^2$. Alors pour tout $\delta \in (0,1)$, il existe une constante $C_\delta > 0$ telle que :*

$$\forall n \in \mathbb{N}, \quad \sup_{(p_1, p_2) \in [0,1], p_2 \leq p_1 \wedge (1-\delta)} \bar{R}_n \leq C_\delta \sqrt{n}.$$

Remarque La borne n'est donc pas complètement uniforme et l'on note que lorsque p_2 est trop grand, l'algorithme a un rappel trop faible pour garantir la bornitude de la suite $((1 - X_n)/\rho_n)_{n \geq 1}$ quand X_n est proche de 0.

Comme nous venons de le voir, la recherche d'un algorithme optimal passera par une nouvelle amélioration. Pour ce faire nous proposons un algorithme dérivé du bandit pénalisé et qui cette fois-ci va pénaliser un bras même lorsqu'il gagne. Ceci permet de gérer le problème rencontré quand X_n est proche de zéro et ainsi d'obtenir un rappel uniforme de l'ordre de \sqrt{n} . Nous appellerons

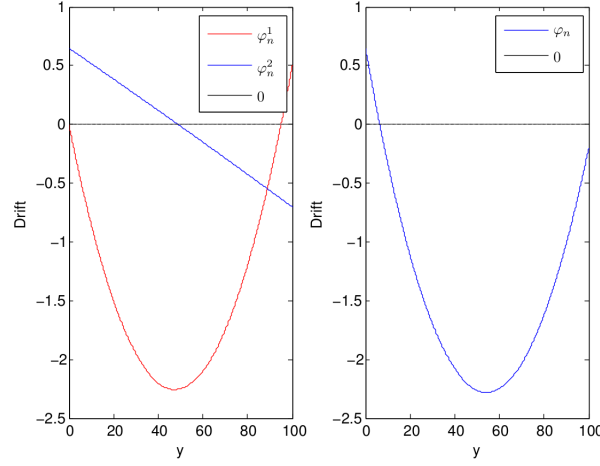


FIGURE 1.1 – Décomposition du drift à gauche et drift global à droite.

NS sur-pénalisé l'algorithme suivant,

$$X_{n+1} = X_n + \begin{cases} \gamma_{n+1}(1 - X_n) - \rho_n \gamma_{n+1} X_n & \text{si le bras 1 est sélectionné et gagne.} \\ -\gamma_{n+1} X_n + \rho_{n+1} \gamma_{n+1} (1 - X_n) & \text{si le bras 2 est sélectionné et gagne.} \\ -\rho_n \gamma_{n+1} X_n & \text{si le bras 1 est sélectionné et perd.} \\ \rho_{n+1} \gamma_{n+1} (1 - X_n) & \text{si le bras 2 est sélectionné et perd.} \end{cases} \quad (1.2.11)$$

THEOREME 1.2.7. Soit $(X_n)_{n \geq 0}$ l'algorithme de bandit sur-pénalisé avec $(\gamma_n, \rho_n)_{n \geq 1}$ tels que $(\gamma_1, \rho_1) \in (0, 1)^2$. Alors, le choix $\sigma = 0$, $\gamma_n = 2.63\rho_n = 0.89/\sqrt{n}$ entraîne la borne suivante

$$\forall n \in \mathbb{N}, \sup_{(p_1, p_2) \in [0, 1], p_2 < p_1} \bar{R}_n \leq 31.1\sqrt{2n}. \quad (1.2.12)$$

Le principe de l'algorithme est d'introduire un nouvel aléa qui pénalise le bras tiré en cas de victoire avec probabilité $1 - \sigma$. Décrit comme (1.2.11), nous avons alors $\sigma = 0$ c'est-à-dire qu'un bras gagnant est toujours pénalisé.

Remarque La version proposée ici de l'algorithme est simplifiée afin de permettre une lecture plus aisée du mécanisme de l'algorithme. Il est à noter que cette borne a été obtenue en effectuant une optimisation de la constante $C_\sigma(\gamma_1, \rho_1)$. Nous avons observé que la valeur de σ optimale semble être proche de 0 et avons ensuite déduit la valeur des paramètres restants par calculs. Cette borne pourrait être améliorée (au prix d'une technicité calculatoire accrue).

Une idée de la preuve

Nous esquissons ici les grandes lignes de la preuve du résultat en donnant notamment le point clé : la montée en puissance.

Nous l'avons dit précédemment, pour obtenir une borne minimax pour le regret il suffit de montrer

que $\sup_{n \geq 1} \mathbb{E}(Y_n) < \infty$. Commençons par reprendre l'expression de Y_n

$$Y_{n+1} = Y_n(1 + \gamma_{n+1}(\epsilon_n - \pi X_n)) - \gamma_{n+1}\kappa(X_n) + \gamma_{n+1}\Delta M_{n+1} \quad (1.2.13)$$

En ajoutant et soustrayant $\pi\gamma_{n+1}Y_n$ nous obtenons

$$Y_{n+1} = Y_n(1 + \gamma_{n+1}(\epsilon_n - \pi)) + \pi\gamma_{n+1}\frac{(1 - X_n)^2}{\rho_n} - \gamma_{n+1}\kappa(X_n) + \gamma_{n+1}\Delta M_{n+1} \quad (1.2.14)$$

Le point clé de cette astuce est le fait que pour un entier n assez grand nous avons $\epsilon_n \leq \pi$. Nous obtenons alors une récurrence de la forme suivante

$$\begin{aligned} \mathbb{E}(Y_{n+1}) &= \mathbb{E}(Y_{n_0}) \prod_{k=n_0}^{n+1} (1 - \gamma_{k+1}(\pi - \epsilon_k)) + \pi \sum_{k=n_0}^{n+1} \gamma_{k+1} \mathbb{E} \left[\frac{(1 - X_k)^2}{\rho_k} \right] \prod_{l=k+1}^{n+1} (1 - \gamma_{l+1}(\pi - \epsilon_l)) \\ &+ \sum_{k=n_0}^{n+1} \gamma_{k+1} \mathbb{E}(\kappa(X_k)) \prod_{l=k+1}^{n+1} (1 - \gamma_{l+1}(\pi - \epsilon_l)) \end{aligned}$$

où $n_0 = \inf_{n \geq 1} \{\epsilon_n \leq \pi\}$. En remarquant que

$$\sum_{k=n_0}^{n+1} \gamma_{k+1} \prod_{l=k+1}^{n+1} (1 - \gamma_{l+1}(\pi - \epsilon_l)) < \infty$$

On se retrouve ramener à montrer que

$$\sup_{n \geq 1} \mathbb{E} \left[\frac{(1 - X_n)^2}{\rho_n} \right] < \infty$$

On peut voir que cette quantité sera plus facile à contrôler et permettra de résoudre les problèmes rencontrés précédemment. En effet, elle est plus petite que Y_n car $X_n \in [0, 1]$ pour tout entier n . C'est pour cela que l'on définit la quantité suivante

$$Z_n^{(r)} = \frac{(1 - X_n)^r}{\rho_n}, \quad r \geq 1.$$

Il reste maintenant deux questions à résoudre :

- Jusqu'à quelle valeur de r va-t-on devoir aller ?
- Comment étudier $Z_n^{(r)}$?

Commençons par remarquer qu'en utilisant la formule du binôme de Newton on obtient

$$Z_{n+1}^{(r)} = Z_n^{(r)}(1 + \gamma_{n+1}(\epsilon_n - r\pi)) + r\pi\gamma_{n+1}Z_n^{(r+1)} + \beta_n^{(r)} + r(1 - X_n)^{r-1}\Delta M_{n+1}$$

On note comme précédemment que l'on choisit un entier n_0 à partir duquel la quantité $\epsilon_n - r\pi$ est négative pour créer un rappel en notant que plus r est grand plus cette quantité est négative. Ceci conduit à la résolution de l'équation suivante :

$$\epsilon_n \leq \epsilon \Leftrightarrow n \geq n_0(\epsilon, \pi, \gamma_1) := \lfloor \frac{1}{4\epsilon^2\pi^2\gamma_1^2} \rfloor + 1$$

Ainsi on déduit que pour tout $n \geq n_0(\epsilon, \pi, \gamma_1)$ on a

$$\epsilon_n - r\pi \leq 0 \Leftrightarrow 2\epsilon\gamma_1^2(r - \epsilon) \leq 1$$

à condition d'avoir $\epsilon \in (0, 1/3)$ l'égalité est vraie pour tout $\gamma_1 \in (0, 1)$ et surtout pour tout entier $r \in \{1, 2\}$. Ainsi, nous arrivons à la conclusion qu'une étude de la suite pour $r = 3$ sera suffisante. Le restant de la preuve consiste alors en une étude fine de la suite $Z_n^{(3)}$ que nous ne détaillerons pas ici car elle beaucoup trop technique pour cette section.

Pour terminer, le lecteur trouvera dans le chapitre 3 un exemple de calibration des divers paramètres entrant en jeu pour lequel une borne explicite est obtenue. Il est à noter que ladite calibration n'est pas optimale et obtenir une meilleure borne est très difficile. Il s'agit en effet d'optimiser une quantité dépendante de 5 paramètres et autrement que numériquement la tâche semble impossible. Quoiqu'il en soit nous obtenons le bon ordre de convergence ainsi qu'une borne minimax en " n ".

1.2.3 Le PMDM du bandit

La première partie consistait en une étude très fine du comportement non-asymptotique des algorithmes NS. Dans cette partie, nous nous intéressons au comportement en temps long de ces algorithmes. Comme nous l'avons vu dans le premier chapitre, un algorithme stochastique correctement renormalisé converge en général vers un processus de diffusion. Dans le cas du bandit, le résultat est surprenant.

PROPOSITION 1.2.4. *Supposons que $\alpha = \beta \leq 1/2$ et $g = \frac{\gamma_1}{\rho_1}$. Alors, Y_n converge en loi vers la mesure invariante d'un processus de Markov agissant sur les fonctions de classe C^1 à support compact dans \mathbb{R}^+ de la manière suivante :*

$$Lf(x) = (1 - p_1 - p_1x)f'(x) + \frac{p_2}{g}x(f(x+g) - f(x))$$

Remarque On trouvera dans le chapitre 2 une version multidimensionnelle du résultat. Par souci de clarté et de pédagogie, nous limitons les énoncés à la version un-dimensionnelle dans la suite.

Ce type de générateur caractérise une classe particulière de processus de Markov. Introduit par Davis en 1983, les PMDM sont des processus qui sont l'essentiel du temps déterministes et dont la trajectoire est ponctuée de sauts à des temps aléatoires et dont la position après le saut est elle-même aléatoire. Les motivations pour l'étude de ces processus sont d'origine diverses. On peut citer par exemple le travail de Malrieu et al. (voir [14]) qui se focalise sur la transmission de données sur internet ou encore le travail de Guérin et al. (voir [49]) sur la modélisation du mouvement de bactéries, en neurobiologie pour le modèle de Hodgkin-Huxley dans l'étude de l'activité neuronale (K. Pakdaman and al. [[102]]), la fiabilité d'automates (F. Dufour and Y. Dutuit voir [45]), en biologie des populations (H.G. Othmer and al. voir [99]).

Plaçons nous sur un ouvert Ω de \mathbb{R}^d dont nous noterons la frontière $\partial\Omega$ et $\bar{\Omega}$ son adhérence. Un PMDM est un processus qui se définit par le biais de trois caractéristiques :

- le flot ϕ qui définit la trajectoire entre les sauts.
- le taux de saut λ qui détermine l'intensité des sauts.
- le noyau de transition Q qui permet de déterminer la position du processus après un saut.

Cette formulation très générale se traduit souvent par le générateur infinitésimal associé. Nous nous placerons en dimension 1 dans ce qui suit afin de permettre au lecteur une approche plus aisée de ces processus. Le générateur d'un PMDM est de la forme :

$$Lf(x) = \underbrace{g(x)f'(x)}_{\text{Flot Déterministe}} + \underbrace{\lambda(x) \int_{\Omega} (f(y) - f(x))Q(x, dy)}_{\text{Saut Aléatoire}}$$

Dans ce cas, on peut préciser la nature du flot qui vérifie l'équation suivante :

$$\begin{aligned}\partial_t \phi(x, t) &= g(x) \partial_x \phi(x, t) \\ \phi(x, 0) &= x\end{aligned}$$

Précisons aussi qu'un PMDM peut aussi être considéré au travers de la chaîne de Markov obtenue en ne considérant que les positions du processus au moment du saut.

Dans le cas qui va nous intéresser la forme du générateur est de la suivante :

$$Lf(x) = (a - bx)f'(x) + cx(f(x + g) - f(x)) \quad (1.2.15)$$

où dans le cas du bandit on a

$$a = 1 - p_1, b = p_1, g = \frac{\gamma_1}{\rho_1}, c = \frac{p_2}{g}$$

et f est une fonction de classe C^1 sur \mathbb{R}^+ .

On peut alors définir le flot associé à ce PMDM. Il est défini par

$$\begin{cases} \partial_t \phi(x, t) &= (a - bx) \partial_x \phi(x, t) \\ \phi(x, 0) &= x \in \mathbb{R}^+ \end{cases}$$

c'est-à-dire

$$\phi(x, t) = \frac{a}{b} + (x - \frac{a}{b})e^{-bt}.$$

Ainsi, ϕ est croissant (resp. décroissant) en temps si $x > \frac{a}{b}$ (resp. $x < \frac{a}{b}$). Ce processus suit donc la trajectoire suivante : le flot est exponentiel entre les sauts et ces derniers sont d'intensité cx et additifs de taille g . Il est à noter que g étant fini et positif, le processus est défini p.s sur \mathbb{R}^+ , positif et n'explose pas en temps fini. La figure 1.2.3 illustre ce comportement.

Existence d'une mesure invariante Il est facile de voir que le processus possède une unique mesure invariante. L'existence est assurée par la fonction de Lyapounov $V(x) = x$ car appliquée au générateur, on obtient la formule suivante

$$LV(x) = a - (b - cg)V(x)$$

Cette égalité assure aussi la non-explosion du système. L'unicité, quant à elle, est assurée par les résultats que nous allons établir (notamment la convergence pour les distances de Wasserstein).

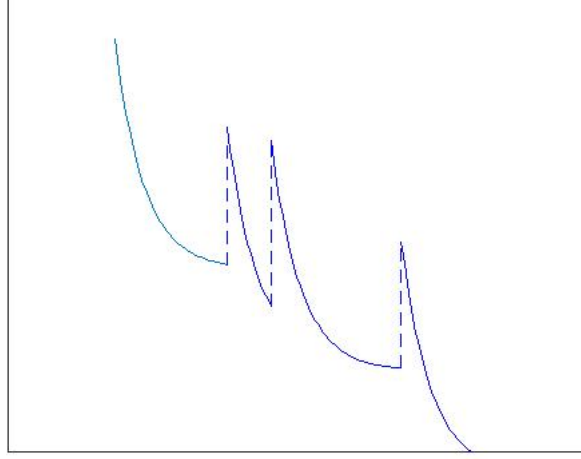


FIGURE 1.2 – Exemple de trajectoire du PMDM du bandit.

Distances de Wasserstein et en Variation Totale

En plus d'être un moyen pour montrer l'unicité de la mesure invariante du PMDM, l'obtention de vitesse de convergence présente un intérêt pratique important. En effet, en étant capable de donner des vitesses précises nous pourrions établir des intervalles de confiances pour les paramètres de succès. Les difficultés sont cependant nombreuses car la mesure invariante est pratiquement inconnue (en dehors de son support) et les PMDM ne sont pas des processus réversibles ce qui rend l'usage de distance telle que la distance L^2 ([51]) impossible. Pour ces raisons, nous travaillerons avec des distances moins sophistiquées mais plus aisées à manier. Dans un premier temps, travaillons avec les distances de Wasserstein.

La distance de Wasserstein d'ordre $p \geq 1$ est définie pour tout couple de mesures de probabilités μ et ν de \mathbb{R}^d par :

$$\mathcal{W}_p(\mu, \nu) = \inf \left\{ \mathbb{E}(|X - Y|^p)^\frac{1}{p} \mid \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu \right\}.$$

Du fait de cette définition, il est facile de voir que trouver une borne pour un couple de mesure suffit car la distance est définie par un inf. On note aussi qu'en contrôlant les moments du processus et en créant un couplage $(X_t, Y_t)_{t \geq 0}$ de sorte que l'ordre entre les deux composantes soit préservé au cours du temps, nous arriverons à borner cette distance. Ceci donne lieu au résultat suivant :

THEOREME 1.2.8. *Soit $p \geq 1$ et notons pour tout $t \geq 0$ $\mu_t := \mathcal{L}(X_t^{\mu_0})$ où $(X_t^{\mu_0})$ est un processus de générateur L avec pour mesure de départ μ_0 (de support inclus dans \mathbb{R}_+^*). Si $p = 1$, nous avons*

$$\left| \int x(\mu_0 - \mu_\infty)(dx) \right| e^{-\pi t} \leq \mathcal{W}_1(\mu_t, \mu_\infty) \leq \mathcal{W}_1(\mu_0, \mu_\infty) e^{-t\pi}$$

et si $p > 1$, il existe une constante γ_p telle que

$$\mathcal{W}_p(\mu_t, \mu_\infty) \leq \gamma_p e^{-\frac{t\pi}{p}}.$$

où $(\gamma_p)_{p \geq 1}$ vérifie la relation de récurrence suivante $\gamma_p^p = \gamma_{p-1}^{p-1}[pa + (1+g)^p]$.

La seconde distance qui va nous intéresser est la distance en Variation Totale. Elle est définie comme suit :

$$\|\mu P_t - \nu P_t\|_{TV} = \inf \mathbb{P}(X \neq Y)$$

où P_t est le semi-groupe du processus de Markov considéré. L'infimum est pris sur tous les couples (X, Y) tels que $\mathcal{L}(X) = \mu$, $\mathcal{L}(Y) = \nu$.

En tirant profit du couplage utilisé pour obtenir les bornes en distances de Wassertein, nous allons essayer de rapprocher les deux trajectoires dans un compact bien choisit pour ensuite essayer de les coller. Ceci donne lieu au résultat suivant :

THEOREME 1.2.9. *Soit μ_0 la distribution de départ du processus dont on supposera qu'elle a des moments de tout ordre. Alors pour tout $\varepsilon > 0$, il existe un réel $C_\varepsilon > 0$ tel que :*

$$\|\mu_0 P_t - \mu_\infty P_t\|_{TV} \leq C_\varepsilon e^{-(\alpha\pi - \varepsilon)t} \quad \text{with } \alpha = \frac{1}{2 + \frac{b\pi}{ac}}.$$

Cette borne semble à la bonne échelle en temps même si les constantes entrant en jeu dans l'exponentielle sont loin d'être optimales.

Idée des preuves

Ces preuves reposent sur des arguments de couplage. Pour le lecteur non familier de ces méthodes nous définissons les termes principaux qui entrent en jeu.

Nous commençons par définir la notion de couplage primordiale dans notre étude.

DEFINITION 1.2.5. Soient μ et ν deux mesures de probabilités. Un couplage de μ et ν est un couple de variables aléatoires (X, Y) suivant respectivement les lois μ et ν .

Pour quantifier la vitesse de convergence pour la distance de Wasserstein, on va se donner deux variables aléatoires dont la dynamique est celle induite par le générateur du PMDM. Du fait de sa définition, obtenir une vitesse de convergence pour la distance de Wasserstein se fait en obtenant une vitesse de convergence pour un couplage quelconque. Ici la propriété centrale est la monotonie stochastique. Autrement dit, si on donne un couplage $(X_t, Y_t)_{t \geq 0}$ tel que $X_0 \geq Y_0$ alors pour tout $t \geq 0$ on a $X_t \geq Y_t$. Le PMDM du bandit vérifie cela car les saut sont positifs et l'on peut ainsi contrôler l'ordre entre les deux marginales du couplage en s'assurant que la plus haute soit la seule à pouvoir effectuer des sauts seule. Cette propriété va jouer un rôle primordiale dans la construction de nos couplage qui vérifieront cette propriété et permettront ainsi l'obtention de vitesse. La figure 1.2.3 illustre cette idée. Mathématiquement, cela se traduit par un générateur infinitésimal qui

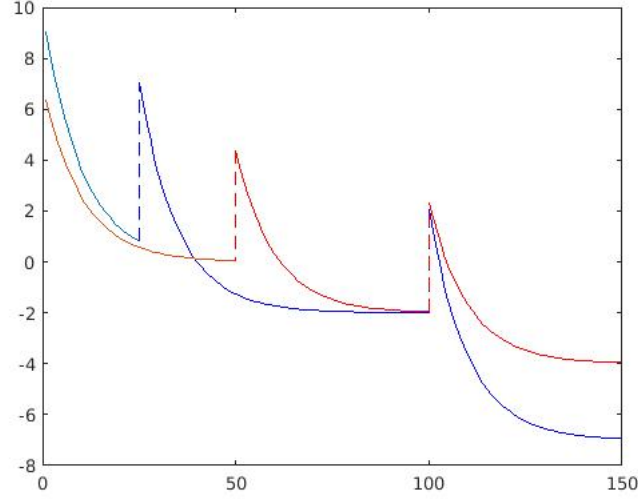


FIGURE 1.3 – Un exemple de trajectoire pour un couplage stochastiquement monotone. La composante la plus haute saute seule ou alors les deux composantes sautent en même temps.

caractérise un tel couplage :

$$\begin{aligned}
 Lf(x, y) = & \underbrace{(a - bx)\partial_x f(x, y) + (a - by)\partial_y f(x, y)}_{\text{flot des trajectoires}} \\
 & + \underbrace{cy(f(x + g, y + g) - f(x, y))}_{\text{saut simultané}} + \underbrace{c(x - y)(f(x + g, y) - f(x, y))}_{\text{saut seule de la composante la plus haute}}
 \end{aligned} \tag{1.2.16}$$

avec une expression symétrique dans le cas $y > x$. La suite de la preuve se résume à remarquer les processus X_t et Y_t vérifient la propriété suivante :

$$\mathbb{E}(|X_t - Y_t|)^p = \mathbb{E}(X_t - Y_t)^p$$

Cette expression permet de ramener l'étude des distances de Wasserstein à celle des moments du générateur définit ici. Il est à noter que la borne obtenue pour ces distances semble être optimale comme le suggère le cas $p = 1$. Dans effet le cas $p = 1$, nous possédons une formulation variationnelle explicite qui nous permet de minorer la distance \mathcal{W}_1 . Plus précisément, nous avons le résultat suivant :

$$\left| \int x(\mu_0 - \mu_\infty)(dx) \right| e^{-\pi t} \leq \mathcal{W}_1(\mu_t, \mu_\infty) \leq \mathcal{W}_1(\mu_0, \mu_\infty) e^{-t\pi}$$

Pour la distance en Variation Totale, le schéma de preuve change. En effet, nous considérons un couplage et nous allons attendre que les deux trajectoires soient suffisamment proches avec une probabilité contrôlée par la borne de Wasserstein obtenue précédemment. Ensuite pour coller les trajectoires, il faut faire sauter une des composantes avec un temps de retard afin qu'elle se retrouve

collée à l'autre. La figure 1.2.3 illustre cela.

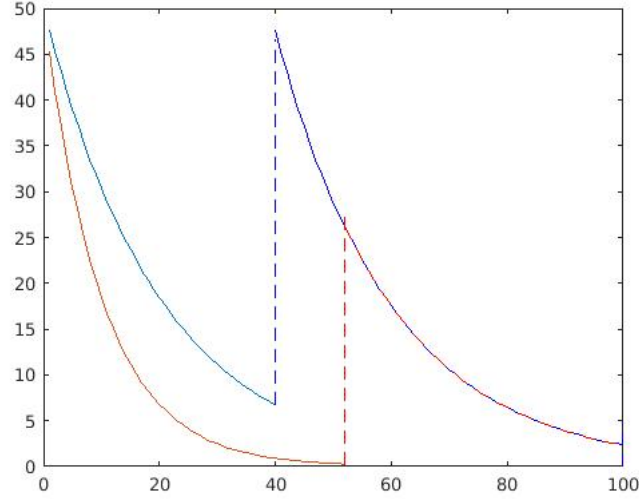


FIGURE 1.4 – La composante rouge saute avec un décalage pour atteindre la composante bleue.

Le travail consiste alors à construire à construire un couplage coalescent (un couplage tel que les deux trajectoires soient égales à partir d'un certain). Plus précisément, il s'agit de rapprocher les composantes au moyen du couplage de Wasserstein pendant un intervalle de temps $[0, t_1]$. Ensuite pendant un intervalle de temps $[t_1, t_2]$, nous essayons de coller les trajectoires une fois qu'elles se trouvent dans un compact dans lequel on peut contrôler de manière fine la probabilité de couplage. C'est ce que donne le lemme suivant.

LEMME 1.2.1. *Soient $\varepsilon > 0$ et $t \geq \frac{1}{b} \ln(1 + \varepsilon)$. Il existe un couplage $(X_t, Y_t)_{t \geq 0}$ dont la dynamique est donnée par (1.2.15) et tel que sur $A_{x_0, \varepsilon}$*

$$\mathbb{P}(X_t = Y_t, t \geq s) \geq \left(1 - \frac{c}{b} x_0 \varepsilon - e^{-\frac{a}{b} cs} - \frac{c\varepsilon}{b}\right) \max(0, 1 - \frac{c}{b} \varepsilon(x_0 + g)),$$

où $A_{x_0, \varepsilon} = \{(x, y) | \frac{a}{b} < x \leq x_0, 0 < x - y \leq \varepsilon\}$.

Il est à noter que la preuve que nous donnons semble loin de fournir une constante optimale mais le résultats est bien meilleur que celui que l'on obtiendrait en utilisant le théorème de Roberts et Rosenthal [15].

1.3 L'algorithme de boule pesante stochastique

Les procédures d'optimisation d'une fonction, dont on a peu (ou pas) d'informations, ont focalisé l'attention d'un grand nombre de travaux. Les premiers remontent à presque 400 ans avec Newton et sa fameuse méthode de recherche des zéros d'une fonction. Les objectifs varient fortement avec la nature des fonctions considérées. On peut penser notamment aux méthodes de descente de gradient (et à ses nombreuses adaptations) où des résultats pertinents sont obtenus dans le cas de fonction ne présentant qu'un seul minimum. Dans le même esprit, on trouvera une multitude de résultats pour des fonctions ayant un seul minimum où l'objectif est de construire une procédure la plus rapide possible alors que les méthodes permettant de traiter le cas de fonctions ayant plusieurs minima sont basées sur une bonne exploration de l'espace. Naturellement, ces différents objectifs amènent à des vitesses de convergence différentes. Ainsi on trouve les cas suivants de vitesse. Tout d'abord, la méthode de Newton est quadratique (elle produit une vitesse en e^{-n^2} où n est le nombre d'itération de la procédure) à condition de l'initialiser de manière adéquate. Les méthodes de descente de gradient sont linéaires (e^{-n}) pour des fonctions fortement convexes (*i.e.* fonctions convexes dont la hessienne est définie positive). Pour des fonctions convexes plus générales (non fortement convexes en particulier), on trouve des bornes en $O(n^{-1})$. Enfin, il existe des méthodes plus rapides en $O(n^{-2})$. Ce sont principalement les méthodes dites d'ordre 2 car elles font appel à deux points : un qui va converger vers le minimum et le second qui permet de gérer les problèmes rencontrés par les méthodes de type descente de gradient. La méthode de Nesterov est l'une des méthodes d'ordre 2 les plus célèbres et l'on sait que dans les résultats qui donnent des bornes inférieures pour la vitesse de convergence (voir [95]), les fonctions construites utilisées sont d'ordre 2. La méthode de Nesterov produit ainsi une vitesse optimale (voir [96]) et reste encore à ce jour un mystère car il n'existe pas (à notre connaissance) d'explication claire et précise pour justifier l'efficacité d'une telle méthode. Cependant, les méthodes de descente ne fonctionnent plus quand plusieurs minima locaux existent.

Un algorithme de descente de gradient ne peut, en effet, en aucun cas distinguer un minimum local d'un minimum global et il faut alors réfléchir à de nouvelles stratégies pour résoudre ce problème. Une solution est d'introduire une composante aléatoire (via un mouvement brownien), c'est l'idée de Kirkpatrick, Gelatt et Vecchi qui, en 1983, proposèrent l'algorithme de recuit simulé. L'idée est de donner à l'algorithme une excitation suffisante afin de ne pas être piégé par un minimum local. Intuitivement, la composante aléatoire crée une excitation qui relance l'algorithme à l'approche d'un minimum local. Malheureusement, cet algorithme est trop lent car il visite beaucoup de zones de l'espace. Il faut alors songer à de nouvelles stratégies.

En effet, Il faut offrir à l'algorithme un moyen d'explorer l'espace et surtout d'apprendre de son passé (à l'instar des algorithmes de bandit). C'est en quelque sorte l'idée de la descente de gradient moyennée introduite par Polyak (1964) où l'on considère un algorithme de descente de gradient avec une moyennisation sur tout le passé de la trajectoire qui offre à l'algorithme une mémoire lui permettant d'apprendre de son passé. Notre travail a pour but d'étudier une version stochastique de cette procédure et de comprendre son fonctionnement tant sur le point théorique que pratique. En effet, nous dégagerons dans ce travail des résultats de convergence pour des fonctions générales, des vitesses pour des fonctions convexes et même un théorème de convergence en loi dans un cas spécifique.

Contenu de la sous-partie

Dans cette partie, notre objectif est de présenter les résultats de l'article *Stochastic Heavy Ball*. Dans cette optique, nous commencerons par rappeler quelques uns des résultats associés au processus de boule pesante afin de poser un cadre naturel pour notre étude. Ensuite, nous présenterons nos résultats qui sont axés sur la convergence d'une version stochastique de l'algorithme de Polyak. Nous verrons que l'on peut garantir la convergence sous des hypothèses raisonnables pour une classe de fonctions large puis nous verrons comment on peut assurer la convergence vers un minimum local. L'obtention de vitesse est en revanche plus complexe mais sera effectuée dans le cas des fonctions fortement convexes avec un résultat fortement dépendant des paramètres du problème. Enfin, nous terminons par un résultat de convergence en loi dans le cadre des mémoires exponentielles.

1.3.1 La descente de gradient

Construire des procédures afin de minimiser une fonction dont on a peu d'informations au préalable est une tâche qui bénéficie d'une littérature très vaste du fait, notamment, des nombreux domaines d'applications de telles procédures. On pourra citer la mécanique des fluides avec notamment l'étude des écoulements diphasiques (écoulement de plusieurs fluides ensemble) qui fait appel à des méthodes de gradient conjugué (voir [84]). On peut aussi penser à toutes les thématiques connexes au *machine learning* où le but est la minimisation d'une fonction de la forme

$$f(\cdot) + \lambda \|\cdot\|_\alpha$$

où f est une fonction convexe et $\alpha \in \{1, 2\}$. On parle de régression Ridge si $\alpha = 2$ et on parle de problème de LASSO pour $\alpha = 1$ voir [119].

L'une des méthodes les plus anciennes et des plus connues est celle de Newton. Appelons f la fonction à minimiser et supposons la deux fois différentiable. La méthode de Newton unidimensionnelle est l'algorithme suivant :

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

Le graphique 1.3.1 décrit la méthode via une représentation graphique. Cette méthode possède un

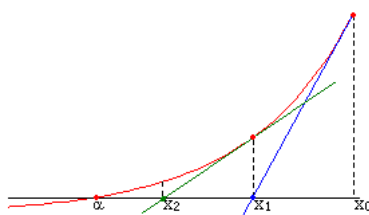


FIGURE 1.5 – Méthode de Newton sur un exemple simple.

inconvenient de taille à savoir la division par la dérivée seconde de la fonction f . Ceci exclut un bon nombre de fonction même s'il est connu que dans un intervalle proche du point de convergence, la vitesse de l'algorithme est géométrique. Pour palier ce problème, on pourrait envisager de remplacer

la dérivée seconde par une constante et c'est ce qui donne lieu à une autre méthode d'optimisation vastement étudiée : *la descente de gradient*.

La descente de gradient est, comme son nom l'indique, un algorithme qui progresse suivant la pente induite par le gradient. Plus concrètement, l'algorithme prend la forme suivante

$$x_{n+1} = x_n - \beta \nabla f(x_n) \quad (1.3.17)$$

où β est une constante que l'on ajuste en fonction de la régularité de la dérivée seconde de f dans un cas idéal. Cette méthode s'interprète comme une minimisation à chaque étape de la meilleure approximation quadratique (obtenue par un développement de Taylor) de la fonction f (voir [97]). On peut facilement comprendre les limites de cet algorithme via un petit calcul que nous détaillons.

$$f(x_{n+1}) - f(x_n) \leq -\beta f'(x_n)^2 + \frac{\beta^2}{2} f''(\xi_n) f'(x_n)^2$$

où $\xi_n \in [x_n, x_{n+1}]$. En supposant f'' majorée uniformément par une constante L , on parvient à l'inégalité suivante

$$f(x_{n+1}) - f(x_n) \leq -(\beta + L \frac{\beta^2}{2}) f'(x_n)^2$$

Une optimisation en β amène au choix suivant

$$\beta = \frac{1}{L}$$

Ces considérations posent des questions sur l'application de cette méthode. En effet, la calibration de β est problématique car le coefficient L n'est pas connu de l'utilisateur de la procédure dans la plupart des cas. On peut néanmoins contourner ce problème en choisissant une méthode d'estimation pour approcher cette valeur ou alors utiliser une borne supérieure. Mais le véritable problème de cette méthode est de ne pas être capable de distinguer un extremum local d'un extremum global. Prenons l'exemple de la fonction définie sur la figure (1.6) qui représente un cas où la fonction possède un minimum local en 1,39 et un minimum global en 4,3. La figure (1.7) illustre les convergences vers ces points en fonction du point d'initialisation (et plus précisément en fonction de la proximité de ce dernier par rapport à un point critique). On comprend ainsi que la méthode est capable de converger vers un point critique mais est incapable de distinguer si c'est un extremum global ou local. Le calcul nous montre que la descente de gradient est une méthode efficace lorsque le gradient prend de grandes valeurs et, a contrario, c'est une méthode inadaptée pour des fonctions dont le gradient prend des petites valeurs. Si on compare avec les algorithmes de la première partie, on voit que les faiblesses du gradient sont semblables à celle de l'algorithme NS non pénalisé. En effet, pour être en mesure de distinguer les extremum locaux et globaux, un bon algorithme doit explorer l'espace afin de ne tomber dans le piège que peut représenter un point critique. Dans cette optique, l'ajout d'une composante aléatoire peut constituer une idée intéressante.

Le gradient stochastique

Dans cette partie, on rappelle des résultats classiques sur l'algorithme du gradient stochastique afin de voir quelles sont les forces et faiblesses de l'algorithme. On considère donc une version

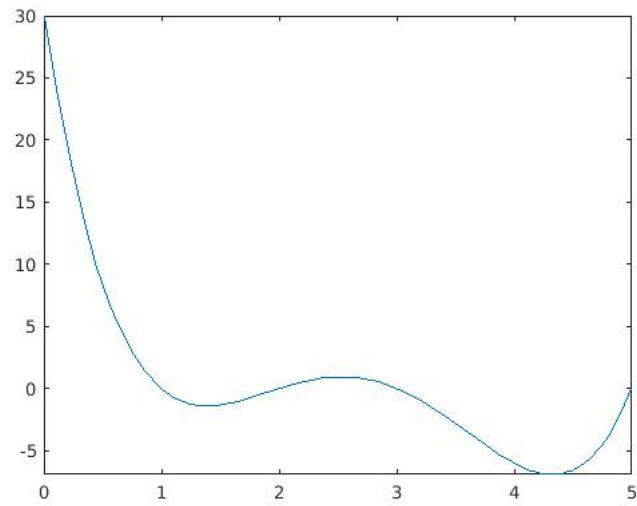


FIGURE 1.6 – Graphique de la fonction $f : x \mapsto x^4 - 11x^3 + 41x^2 - 61x + 30$.

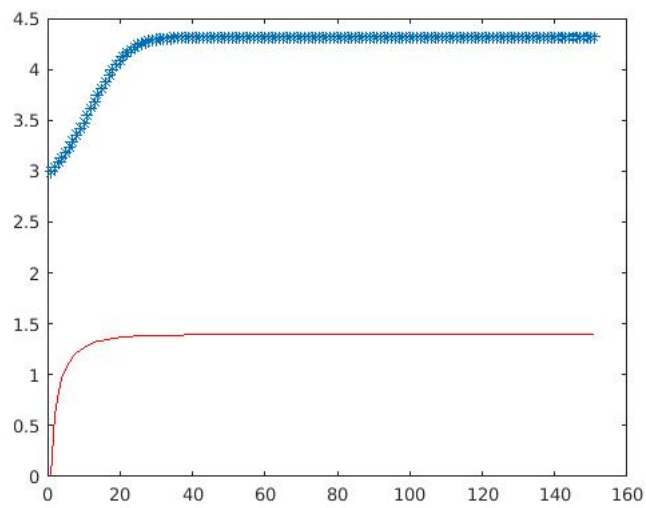


FIGURE 1.7 – Convergence du gradient pour la fonction $f : x \mapsto x^4 - 11x^3 + 41x^2 - 61x + 30$.

stochastique de l'algorithme (1.3.17) définie comme suit

$$X_{n+1} = X_n - \gamma_{n+1} \nabla V(X_n) + \gamma_{n+1} \Delta M_{n+1} \quad (1.3.18)$$

où V est une fonction continue et $(\Delta M_n)_{n \geq 1}$ est une suite d'incrément de martingale. Nous supposons que

$$\gamma_n = \gamma n^{-\beta}, \quad \text{avec } \beta \in (1/2, 1]. \quad (1.3.19)$$

Cette formulation est une méthode de gradient où l'on a accès seulement à des évaluations bruitées et non biaisées du gradient de la fonction à minimiser, ici la fonction V . Ainsi, sous de bonnes hypothèses, on peut s'attendre à la convergence de l'algorithme comme nous l'avons vu avec le théorème de Robbins-Siegmund. Ceci donne lieu au résultat suivant.

THEOREME 1.3.10. *Soit $(X_n)_{n \geq 1}$ défini par (1.3.18) et $(\gamma_n)_{n \geq 1}$ vérifiant (1.3.19). Sous les hypothèses suivantes :*

- V est de classe C^2 , $\lim_{|x| \rightarrow \infty} V(x) = \infty$, $|\nabla V|^2 \leq C(1 + V)$.
- Pour tout entier n , $\mathbb{E}(|\Delta M_{n+1}|^2 | \mathcal{F}_n) \leq C(1 + V(X_n))$.

Nous avons

$$\sup_{n \geq 1} \mathbb{E}(V(X_n)) < \infty$$

et

$$V(X_n) \text{ converge p.s vers } V_\infty$$

Remarque On notera que ce résultat n'est autre qu'une application du lemme de Robbins-Siegmund. Il garantit donc la convergence de la procédure sans toutefois donner une indication précise de la limite hormis le fait que c'est un point critique (minimum ou maximum ? local ou global ?). De plus, nous n'avons aucune idée de la vitesse de cette convergence.

Ainsi, en ajoutant de la convexité à la fonction V on peut obtenir à la fois la convergence et une idée de la vitesse. On définit les hypothèses suivantes.

(H1) V est convexe, de classe C^1 et il existe $L > 0$ tel que pour tout $x, y \in \mathbb{R}^d$ on a

$$V(x) - V(y) \leq \langle \nabla V(y), x - y \rangle + \frac{L}{2} |x - y|^2$$

(H2) Il existe $\alpha > 0$ tel que pour tout $x, y \in \mathbb{R}^d$ on a

$$V(x) - V(y) \geq \langle \nabla V(y), x - y \rangle + \frac{\alpha}{2} |x - y|^2$$

(H3) Pour tout entier n on a, $\mathbb{E}(|\Delta M_{n+1}|^2 | \mathcal{F}_n) \leq C(1 + V(X_n))$.

Nous présentons le résultat suivante extrait de [10] et qui donne une vitesse pour les fonctions fortement convexes.

THEOREME 1.3.11. *Supposons que $(X_n)_{n \geq 1}$ soit définie par (1.3.18) et $(\gamma_n)_{n \geq 1}$ vérifie (1.3.19). Si de plus, les hypothèses (H1), (H2) et (H3) sont satisfaites alors*

$$\mathbb{E}(|X_n - x^*|^2) \leq C(L, \alpha, \gamma) \gamma_n$$

où $C(L, \alpha, \gamma)$ est une constante dépendante de L, α et γ .

Remarques Le calibrage de la constante ne peut être effectué dans la pratique car la fonction V est inconnue. Les vitesses obtenues sont d'autant meilleurs que β est proche de 1. Cependant (et nous rencontrerons cette situation dans l'article *Stochastic Heavy Ball*), ce pas se traite à part des autres pour des raisons techniques. Il est à noter qu'un résultat existe pour les fonctions non fortement convexe (voir [10]). On trouvera aussi dans [10] divers cas dont une application de la méthode de moyennisation de Ruppert et Polyak. Naturellement, on peut se poser la question de l'optimalité de cette borne. Les travaux de Yudin et Nemirovski (voir [95]) ont prouvé que la vitesse optimales dans le cadre stochastique pour une fonction fortement convexe est de l'ordre de $O(1/n)$ et $O(1/\sqrt{n})$ pour les fonctions convexes. Cette vitesse est donc atteignable mais l'on doit au préalable connaître certains paramètres du problèmes.

Les algorithmes exploratifs

Parmi les autres solutions proposées pour déterminer le minimum d'une fonction présentant plusieurs extrema locaux, l'une des premières est celle qui consiste à ajouter une composante brownienne à la descente de gradient. Introduit par Kirkpatrick, Gelatt et Vecchi en 1983, l'algorithme du recuit simulé est décrit comme suit

$$dX_t = -\nabla f(X_t)dt + \sqrt{\sigma_t}dB_t$$

Ici le point important est le choix de σ_t comme le montre Hajek (voir [61]). En effet, pour obtenir un algorithme pertinent il faut choisir

$$\sigma_t = \frac{K}{\log(t)}, \quad \text{avec } K > 0$$

Cette équation propose une procédure efficace qui converge bien vers le minimum global de f . Cependant, cet algorithme est très lent car il passe beaucoup de temps à explorer certaines zones avant de finir par les quitter. L'application d'une fonction σ logarithmique est donc dans la pratique impossible et c'est pour cela que des fonction géométriques (voir [39]) ou encore des fonctions adaptatives (voir [59]) sont préférées.

L'idée du recuit simulé est donc de donner de l'inertie instantanée à l'algorithme mais comme nous l'avons dit la vitesse pâtit de la trop longue exploration. Il faudrait donc un algorithme qui soit capable d'explorer et d'exploiter les informations glanées au cours des trajets (encore une fois l'équilibre entre exploration et exploitation est au coeur de la problématique). C'est l'idée qui motive les travaux de Polyak (voir [109]) en proposant l'algorithme de descente de gradient à mémoire. L'idée de ce procédé est de ne plus effectuer une descente selon la position courante mais sur une moyennisation de la trajectoire sur tout son passé. Pour ce faire, nous introduisons deux fonctions h et k positives et croissantes et considérons l'algorithme suivant

$$dX_t = -\frac{1}{k(t)} \left(\int_0^t h(s) \nabla f(X_s) ds \right) dt \quad (1.3.20)$$

Cet algorithme a été étudié dans des cadres aussi bien déterministes que stochastiques. Dans le cadre de cette thèse nous nous focaliserons sur le cas où $h = \dot{k}$.

1.3.2 Descente de gradient à mémoire : Panorama des résultats connus

Cadre déterministe

Parmi les travaux ayant inspiré les résultats de cette thèse, on peut commencer par citer les travaux de [35] et [36] qui s'attachent à étudier avec une équation différentielle du second ordre de la forme :

$$\ddot{x} + a(t)\dot{x} + \nabla f(x) = 0 \quad (1.3.21)$$

La fonction a joue évidemment un rôle important dans la dynamique d'un processus satisfaisant (1.3.21). Parmi les cas intéressants, on peut noter celui des fonctions du type $a(t) = \frac{c}{t}$, $c \geq 1$ qui donnent à (1.3.21) une structure d'équation de Bessel. Plus précisément, si on pose $f(x) = \frac{x^2}{2}$ et $a(t) = \frac{1}{t}$ on peut facilement voir que les solutions de (1.3.21) sont de la forme

$$x(t) \simeq C \cdot \frac{1}{\sqrt{t}} \cos(t - \frac{\pi}{4}), \quad t \rightarrow \infty$$

Le système est alors oscillant avec un amortissement autour du minimum de f . Un autre cas pertinent est celui où $a(t) = a > 0$. L'algorithme modélise alors le modèle de boule pesante avec frottements où l'on considère une boule (soumise à une force de rappel $-\nabla f(x)$ et une force de frottement $-a\dot{x}$) évoluant sur un graphe à la recherche du minimum d'une fonction. Ici l'algorithme subit un amortissement constant et n'est ainsi pas un algorithme de descente mais un algorithme qui peut explorer toutes les zones de l'espace via une inertie portée par $-\nabla f(x_t)$ au second ordre.

Les auteurs se sont principalement intéressés aux conditions pour garantir la convergence d'une part. D'autre part, des conditions sont établies pour assurer la convergence vers un minimum de la fonction en question. Il est notamment attendu que le choix de la fonction a joue un rôle prépondérant. C'est ce que met en avant la proposition suivante :

PROPOSITION 1.3.5. *Si on note $V(t) = f(x_t) + \frac{|\dot{x}_t|^2}{2}$ alors $\dot{V}(t) = -a(t)|\dot{x}_t|^2$. En particulier, les solutions de (1.3.21) sont définies sur \mathbb{R}^+ et restent bornées. De plus,*

$$\forall t > 0 \quad V(t) - \min f \geq (V(0) - \min f) \exp(-2 \int_0^t a(s) ds)$$

On déduit de cela que si $\int_0^t a(s) ds < \infty$ alors les trajectoires ne peuvent converger.

Cette proposition limite la forme que peut prendre la fonction a . Pour aller plus loin, il suffit de considérer le cas où $f = 0$. En effet les solutions peuvent être calculées exactement

$$x_t = x_0 + \dot{x}_0 \int_0^t e^{-\int_0^s a(s) ds} dt$$

Ce qui entraîne une garantie de convergence sous la condition

$$\int_0^t e^{-\int_0^s a(s) ds} dt < \infty$$

Ces résultats vont poser naturellement un cadre à l'étude effectuée par [52] puisque dans ce cas, $a(t) = c/t$ avec $c > 1$ ou $a(t) = a$ permettent de garantir ces conditions. Via un changement de variable astucieux cet algorithme est relié aux méthodes de Nesterov et de boule pesante. Nous détaillons ici ces connexions de manière explicite avant de présenter les résultats de [52]. Tout d'abord, nous commençons par expliciter le changement de variable permettant de passer de (1.3.20) à (1.3.21). Si $(x_t)_{t \geq 0}$ est solution de (1.3.20) alors $(\tilde{x}_t)_{t \geq 0}$ est solution de (1.3.21) avec

$$\tilde{x}_s = x_{\tau(s)} \quad \text{et} \quad \dot{\tau}(s) = \sqrt{(k \cdot h^{-1})(\tau(s))} \quad \text{où} \quad a(s) = \frac{\dot{k}h + \dot{h}k}{2h^{3/2}k^{1/2}} \circ \tau(s)$$

En particulier, les résultats de [35] et [36] font ressortir deux types de fonctions k qui présentent un intérêt à être étudié. D'une part, on peut choisir $a_t = r$ avec $r > 0$ ce qui revient à choisir $k(t) = r^2 e^{r^2 t}$ et on a alors $\tau(s) = r\sqrt{s}$. L'algorithme trouve une interprétation physique en s'apparentant à l'algorithme de boule pesante. Cet algorithme représente une boule qui se déplace le long du graphe d'une fonction en étant soumise aux seules forces de gravité et de frottement. Les résultats sont nombreux pour cette méthode et nous renvoyons le lecteur à [56] pour une étude précise pour des fonctions convexes dans le cas discret déterministe. L'autre choix que l'on peut s'autoriser est de prendre $a_t = \frac{r}{t}$ avec $r > 0$ ce qui revient à choisir $k(t) = t^{\frac{r+1}{2}}$ et on a alors $\tau(s) = \frac{s^2}{2r-1}$. L'algorithme est alors apparenté à la fameuse descente de gradient accélérée de Nesterov (1983) décrite par les itérations suivantes :

$$\begin{cases} x_{n+1} = y_n - \beta \nabla f(x_n) \\ y_{n+1} = x_{n+1} + \frac{n}{n+r}(x_{n+1} - x_n) \end{cases} \quad (1.3.22)$$

pù β et r sont des réels positifs. En utilisant un développement de Taylor à l'ordre 2 il est possible de relier (1.3.22) à l'équation différentielle suivante :

$$\ddot{x} + \frac{r}{t}\dot{x} + \nabla f(x) = 0 \quad (1.3.23)$$

Cette connection est intéressante et le travail [114] s'attache à comprendre le rôle de la constante r et d'obtenir des vitesses de convergence pour de telles équations différentielles. En particulier, il en ressort un phénomène particulier lorsque l'on travaille avec (1.3.23). En effet, l'algorithme de Nesterov d'origine est écrit avec $r = 3$ et il se trouve que cette constante est un seuil à partir duquel l'algorithme converge (voir [114]). Tout d'abord, la convergence de l'algorithme se fait à une vitesse $O(1/t^2)$ (voir [114]) et $O(1/n^2)$ (voir [96]) pour tout entier $r \geq 3$ et on observe une détérioration de cette vitesse pour des valeurs de r inférieures à 3. Il existe peu de travaux dans la littérature qui donne une explication quant à l'efficacité de la méthode de Nesterov. On pourra néanmoins citer [2] qui propose une interprétation particulièrement intéressante en mettant en avant que le schéma de Nesterov n'est autre qu'un moyen de gérer les problèmes rencontrés par la descente de gradient en présence de petites valeurs du gradient. Ici, cette constante 3 semble ouvrir une porte quant à une explication via des équations différentielles ce qui constituerait un pont des plus intéressants.

Cadre stochastique

Le cadre de notre travail s'inscrit dans la suite logique de l'article [52] où les auteurs se sont attelés à la compréhension de l'équation suivante :

$$dX_t = -\frac{1}{k(t)} \left(\int_0^t \dot{k}(s) \nabla f(X_s) ds \right) dt + \sigma(X_t) dB_t \quad (1.3.24)$$

Cet algorithme a la particularité de ne pas être markovien (il a de la mémoire) mais il est markovien d'ordre 2. En effet, en posant

$$Y_t = \frac{1}{k(t)} \left(\int_0^t \dot{k}(s) \nabla f(X_s) ds \right) dt,$$

on obtient un processus de Markov en considérant le système suivant :

$$\begin{cases} dX_t &= \sigma(X_t) dB_t - Y_t dt \\ dY_t &= r(t)(\nabla U(X_t) - Y_t) dt, \end{cases}$$

où $r(t) = \frac{\dot{k}(t)}{k(t)}$. Le processus $(X_t, Y_t)_{t \geq 0}$ est alors markovien inhomogène en temps ayant pour générateur infinitésimal :

$$\mathcal{A}f(x, y, t) = -\langle y, \partial_x f \rangle + r(t) \langle U(x) - y, \partial_y f \rangle + \frac{1}{2} \text{Tr}(\sigma(x)^* D_x^2 f \sigma(x)) + \partial_t f$$

avec $f : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^+ \mapsto \mathbb{R}$ une fonction de classe C^2 . Cet algorithme est à rapprocher des équations de Fokker-Planck cinétique où on étudie l'évolution d'un couple position/vitesse $(X_t, V_t)_{t \geq 0}$ solution du système suivant :

$$\begin{cases} dX_t &= V_t dt \\ dV_t &= -(\nabla U(X_t) - V_t) dt + \sigma dB_t, \end{cases}$$

On note cependant une différence de taille puisque le bruit est sur la seconde composante alors que le gradient moyenné a lui un bruit sur la première (pour des raisons techniques essentiellement). L'autre différence est que dans l'étude de l'équation de Fokker-Planck cinétiques, on connaît explicitement la mesure invariante de la diffusion (voir par exemple [121]), alors que pour le gradient moyenné ce n'est pas le cas.

L'étude de [52] porte alors sur le comportement en temps long de l'algorithme. Tout d'abord, l'existence de solutions s'obtient en exhibant une fonction de Lyapounov. On a le résultat suivant tiré de [53].

PROPOSITION 1.3.6. *Supposons que $\text{Tr}(\sigma(x)^* D^2 U \sigma(x)) \leq C U(x)$, alors il existe des solutions (fortes) pour (1.3.24). De plus, si $\mathbb{E}(U(X_0) + |Y_0|^2) < \infty$ alors pour tout $T > 0$, on a $\sup_{0 \leq t \leq T} \mathbb{E}(U(X_t) + |Y_t|^2) < \infty$.*

Les auteurs obtiennent des résultats sur les mesures limites lorsqu'elles existent. D'une part, il y a un travail sur l'hypoellipticité (i.e la régularité de la mesure limite) et ensuite un travail sur la caractérisation de cette dernière. Deux types de fonctions r , appelée *mémoire*, sont à l'étude.

Tout d'abord, la mémoire dite polynomiale correspond au choix $k(t) = \exp(rt)$ avec $r > 0$ et donc $r(t) = r$. Sous les hypothèses suivantes :

- σ et f sont de classes C^∞ et σ est uniformément elliptique.
- $\dim(x \in \mathbb{R}^d | \det(D^2 f(x) = 0)) \leq d - 1$

On obtient alors le résultat suivant :

THEOREME 1.3.12. *Le semi-groupe P_t associé à (1.3.24) admet une densité $p_t(z, \cdot)$. De plus, $z' \mapsto p_t(z, z')$ est de classe C^∞ sur $\mathbb{R}^d \times \mathbb{R}^d$ et $z \mapsto P_t(z, \cdot)$ est fortement fellerien. De plus, si $\liminf_{x \rightarrow \infty} \frac{f(x)}{x} = \infty$, alors pour tout $z \in \mathbb{R}^d$ le support de $z \mapsto P_t(z, \cdot)$ est $\mathbb{R}^d \times \mathbb{R}^d$. Par conséquent, il existe au plus une seule mesure invariante ν_r pour $(P_t)_{t \geq 0}$.*

Si on suppose l'existence de ν_r alors cette mesure admet une densité p_r qui est l'unique solution positive de l'équation suivante

$$\langle y, \nabla_x p_r \rangle + \frac{1}{2} \text{Tr}(\sigma^* D_x^2 p_r \sigma) + r \langle y - \nabla f(x), \nabla_y p_r + p_r \rangle = 0$$

et elle vérifie $\int_{\mathbb{R}^d \times \mathbb{R}^d} p_r(x, y) dx dy = 1$.

Si ce théorème se focalise sur la régularité et l'existence de la mesure invariante, d'autres résultats précisent les comportements possibles en fonction des mémoires. Si dans le cas où $\lim_{t \rightarrow \infty} r(t) > 0$ on peut conclure quant à l'existence d'une mesure invariante, les problèmes commencent quant on s'attaque à la mémoire dite polynomiale i.e $r(t) = \frac{r}{t}$ avec $r > 0$. C'est ce qu'illustre la proposition suivante

PROPOSITION 1.3.7. *Supposons que $|\nabla f|^2 \leq C(1 + f)$ et qu'il existe une constante $\lambda > 0$ tel que $\text{Tr}(\sigma^* D^2 f \sigma)(x) \geq \lambda$. Alors si $r(t) \rightarrow 0$ et que l'on peut trouver $t' > 0$ tel que $r'(t) + 2r^2(t) \geq 0$ pour tout $t \geq t'$, on a alors pour tout $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$,*

$$\limsup_{t \rightarrow \infty} r(t) \mathbb{E}_x(|X_t|^2) > 0$$

Il est possible de rendre ce résultat plus précis en spécifiant une fonction f et dans le cas de la dimension 1. On retrouve alors un théorème central limite dans l'esprit de celui énoncé pour les algorithmes stochastiques.

THEOREME 1.3.13. *Supposons que $d = 1$ et $f(x) = \frac{x^2}{2}$ et $k(t) = (1 + t)^r$ avec $r > 1/2$. Alors*

$$\mathbb{E}_x(|X_t|^2) \sim \frac{t}{2r+1} \quad \text{et} \quad \lim_{t \rightarrow \infty} \mathbb{E}_x(|Y_t|^2) = \frac{r}{2r+1}$$

$$\left(\sqrt{\frac{2r+1}{t}} X_t, \sqrt{\frac{2r+1}{r}} Y_t \right) \Rightarrow^{\mathcal{L}} \mathcal{N}(0, I_2) \quad p.s \quad t \rightarrow \infty$$

Divers cas se présentent et on pourrait résumer les comportements en fonction de la mémoire comme suit. En notant, $r_\infty = \lim_{t \rightarrow \infty} r(t)$, nous avons :

- $r_\infty \in (0, \infty]$: le système est stable et on peut donner des résultats de convergence vers une mesure invariante. On peut même rapprocher le comportement de celui de l'équation différentielle suivante : $dZ_t = -\nabla U(Z_t)dt + \sigma(Z_t)dB_t$.
- $r_\infty = 0$: le système est instable car il a trop de mémoire. On se dit alors que la mémoire polynomiale est un mauvais choix à moins de choisir σ tendant lui aussi vers 0 (quand t tends vers l'infini) suffisamment vite.

1.3.3 Présentation de nos résultats

L'objectif de cette sous-section est de présenter nos principaux résultats. Nous commencerons par poser un cadre de travail en accord avec les résultats déjà existants. Ensuite, nous établirons la convergence de l'algorithme en présence de plusieurs minima locaux et nous obtiendrons aussi des vitesses dans le cas favorable des fonctions fortement convexes. L'étude se terminera par un résultat de convergence en loi.

Définitions et premiers résultats

Le cadre défini par [52] permet naturellement de définir un algorithme stochastique associé à (1.3.20) dans le cas où $h = \dot{k}$. On considère ainsi l'algorithme stochastique défini comme suit, $(X_0, Y_0) \in \mathbb{R}^{2d}$ et

$$\begin{cases} X_{n+1} = X_n - \gamma_{n+1} Y_n \\ Y_{n+1} = Y_n + \gamma_{n+1} r_n (\nabla f(X_n) - Y_n) + \gamma_{n+1} r_n \Delta M_{n+1}, \end{cases} \quad (1.3.25)$$

où nous noterons $(\mathcal{F}_n)_{n \geq 0}$ la filtration naturelle associée à la suite $(X_n, Y_n)_{n \geq 0}$. Au vu des résultats de [35] et [36], il est naturel de considérer deux types de mémoire : les mémoire polynomiale et exponentielle. Comme nous l'avons vu dans la partie sur les algorithmes stochastiques la suite d'incrément de martingales $(\Delta M_n)_{n \geq 0}$ représente l'écart entre la vision bruitée du gradient et sa véritable valeur i.e

$$\Delta M_{n+1} = \nabla f(X_n) - \nabla f(X_n, Z_n)$$

où $(Z_n)_{n \geq 0}$ est une suite de variables aléatoires i.i.d. En particulier, nous avons

$$\mathbb{E}(f(X_n, Z_n) | \mathcal{F}_n) = f(X_n).$$

Ici, le bruit se situe sur la seconde composante alors que dans [52] il est sur la première. Ceci se justifie par la formalisation d'un algorithme stochastique, l'aléa est sur la même composante que la fonction d'intérêt.

La suite $(\gamma_n)_{n \geq 0}$ est le pas de l'algorithme et nous définissons le temps de l'algorithme par

$$\Gamma_n = \sum_{k=1}^n \gamma_k$$

de sorte que

$$\Gamma_n \rightarrow \infty$$

Enfin, la suite $(r_n)_{n \geq 0}$ joue le rôle de la fonction r vue dans le cas continu et est définie par

$$r_n = \frac{\dot{k}(\Gamma_n)}{k(\Gamma_n)}$$

En particulier, si nous choisissons $k(t) = e^{rt}$ on parle de mémoire **exponentielle** et on a

$$r_n = r > 0.$$

Si on choisit $k(t) = t^r$, on parle de mémoire **polynomiale** et on a

$$r_n = \frac{r}{\Gamma_n}.$$

Hypothèses

Dans ce qui suit, $\|\cdot\|$ (resp. $\|\cdot\|_F$) fera référence à la norme euclidienne sur \mathbb{R}^d (resp. la norme de Frobenius sur $\mathcal{M}_d(\mathbb{R})$). Pour toute matrice $A \in \mathcal{M}_d(\mathbb{R})$, on notera $\|A\|_\infty := \sup_{i,j} |a_{i,j}|$ où $a_{i,j}$ fait référence aux coefficients de la matrices A . Il convient de distinguer les hypothèses qui concernent la fonction f , de celles qui concernent le bruit $(\Delta M_n)_{n \geq 0}$ ou la suite de pas de l'algorithme.

Commençons par les hypothèse relatives à la fonction f .

- (\mathbf{H}_s) : la fonction f est de classe $C^2(\mathbb{R}^d, \mathbb{R})$ et vérifie

$$\lim_{|x| \rightarrow \infty} f(x) = \infty, \quad \|D^2 f\|_\infty < \infty, \quad \|\nabla f\|^2 \leq c_f f$$

Cette hypothèse est relativement faible puisque l'on demande seulement à f d'être deux fois différentiable, coercive et à croissance au plus quadratique. Ces hypothèses sont à rapprocher de celles usuellement utilisées pour un algorithme stochastique. Nous attirons l'attention du lecteur sur le fait qu'aucune forme de convexité n'apparaît ici.

- $(\mathbf{H}_{\mathbf{SC}}(\alpha))$: f est convexe avec $\alpha := \inf_{x \in \mathbb{R}^d} Sp(D^2 f(x)) > 0$ et $D^2 f$ est lipschitzienne. En particulier, f est α -fortement convexe et vérifie donc

$$\forall (x, y) \in \mathbb{R}^{2d}, \quad f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle + \frac{\alpha}{2} \|x - y\|^2$$

Évidemment, cette hypothèse est la plus favorable dans le cas d'un problème d'optimisation convexe et elle signifie que la hessienne de f est minorée par α . Supposer $D^2 f$ lipschitzienne est utile pour garantir la convergence vers un minimum local de la procédure.

Les hypothèses sur le bruit sont essentiellement à mettre en lien avec le contrôle des moments.

- $(\mathbf{H}_{\sigma, \mathbf{p}})$: Pour tout entier n et $p \geq 1$ on a

$$\mathbb{E}(\|\Delta M_{n+1}\|^{2p} | \mathcal{F}_n) \leq \sigma^2 (1 + f(X_n))^p$$

C'est une hypothèse standard dans le cas $p = 2$ pour un algorithme stochastique. A noter ici que l'on exigera plus en demandant à cette propriété vraie pour tout entier p afin d'obtenir des vitesses de convergence explicites.

- $(\mathbf{H}_{\mathbf{Gauss}, \sigma})$: Pour tout entier n , la transformée de Laplace du bruit vérifie

$$\forall t \geq 0, \quad \mathbb{E}(\|\exp(t \Delta M_{n+1})\|^{2p} | \mathcal{F}_n) \leq \exp\left(\frac{t^2 \sigma^2}{2}\right).$$

Cette hypothèse traduit le fait que l'on supposera le bruit sous-gaussien. En particulier, $(\mathbf{H}_{\mathbf{Gauss}, \sigma})$ implique $(\mathbf{H}_{\sigma, \mathbf{p}})$. Si elle est restrictive cette hypothèse est cruciale pour obtenir un résultat de convergence p.s dans le cas non convexe.

- $(\mathbf{H}_{\mathcal{E}})$: Pour tout entier n , le bruit vérifie

$$\forall v \in S_{\mathbb{R}^d}^1, \quad \mathbb{E}(|\langle \Delta M_{n+1}, v \rangle| | X_n, Y_n) \geq c_v > 0.$$

Cette hypothèse trouve son utilité au moment d'établir la convergence vers un minimum de f dans le cas non convexe. Elle est à rapprocher des propriétés d'uniforme ellipticité que l'on trouve dans le cas continu et signifie que si on projette le bruit dans n'importe quelle direction, il a une espérance strictement positive.

Les hypothèses sur la pas vont jouer logiquement sur la mémoire r_n .

— $(\mathbf{H}_\beta^\gamma)$: la suite $(\gamma_n)_{n \geq 0}$ vérifie

$$\forall n \in \mathbb{N} \quad \gamma_n = \frac{\gamma}{n^\beta} \quad \text{avec} \quad \beta \in (0, 1],$$

ce qui implique

$$\forall \beta \in (0, 1) \quad \Gamma_n \sim \frac{\gamma}{1-\beta} n^{1-\beta} \quad \text{alors que} \quad \Gamma_n \sim \gamma \log n \quad \text{si} \quad \beta = 1.$$

La mémoire sera donc soit polynomiale i.e $r_n = \frac{r}{\Gamma_n}$ soit exponentielle i.e $r_n = r$ avec $r > 0$. Ainsi sous l'hypothèse $(\mathbf{H}_\beta^\gamma)$, on a pour les deux mémoires

$$\sum_{n \rightarrow \infty} \gamma_n r_n = \infty$$

et

$$\sum_{n \rightarrow \infty} \gamma_n^2 r_n < \infty$$

Nous définissons l'ensemble d'hypothèses suivant :

— (\mathbf{H}_r) : la suite $(r_n)_{n \geq 0}$ vérifie

$$\sum_{n \rightarrow \infty} \gamma_n r_n = \infty \quad \text{et} \quad \sum_{n \rightarrow \infty} \gamma_n^2 r_n < \infty \quad \text{et} \quad \limsup_{n \rightarrow \infty} \frac{1}{2\gamma_{n+1}} \left(\frac{1}{r_n} - \frac{1}{r_{n+1}} \right) =: c_r < 1$$

Dans le cas exponentiel, on a $c_r = 0$ alors que dans le cas polynomial on a $c_r = \frac{1}{2r}$ et il faut alors choisir $r > 1/2$. Nous noterons par la suite $r_\infty = \lim_{n \rightarrow \infty} r_n$ qui vaut soit 0 (mémoire polynomiale) soit r (mémoire exponentielle).

Un premier résultat important

Avant de s'attaquer aux résultats de convergence, nous présentons un résultat de contrôle dans l'esprit du théorème de Robbins-Siegmund. Le contrôle de type Lyapounov s'obtient en considérant la fonction $(x, y) \mapsto V_n(x, y)$ dépendante de deux paramètres $a, b > 0$. Elle est définie par

$$V_n(x, y) = (a + br_{n-1})f(x) + \frac{a}{2r_{n-1}}|y|^2 - b\langle \nabla f(x), y \rangle,$$

On obtient en particulier le résultat suivant

LEMME 1.3.2. *Supposons les hypothèses $(\mathbf{H}_{\sigma, \mathbf{p}})$ et (\mathbf{H}_s) vérifiées. Si de plus, $c_r < 1$ alors pour tout couple $(a, b) \in \mathbb{R}^+$ tel que*

$$\frac{a}{b} > \left(\frac{1}{2} \vee \frac{\|D^2 f\|_\infty}{1 - c_r} \vee r_\infty(c_f - 1) \right),$$

il existe

— (i) une constante $C_1 > 0$ et un entier n_0 tels que pour tout $n \geq n_0$, on a

$$\forall x, y \in \mathbb{R}^d; \quad V_n(x, y) \geq C_1 \left[f(x) + \frac{|y|^2}{2r_{n-1}} \right]$$

— (ii) des constantes positives C_2, C_3 et $c_{a,b}$ telles que

$$\mathbb{E}(V_{n+1}(X_{n+1}, Y_{n+1}) | \mathcal{F}_n) \leq V_n(X_n, Y_n)(1 + C_2 \gamma_{n+1}^2 r_n) - c_{a,b} \gamma_{n+1} |Y_n|^2 - b \gamma_{n+1}^2 r_n |\nabla f(X_n)| + C_3 \gamma_{n+1}^2 r_n$$

Remarques Cette fonction est proche dans sa structure de celle utilisée pour les équations de Fokker-Planck cinétique (voir [121]). Les deux termes $f(x) + |y|^2/2$ génèrent des termes de la forme $a - |y|^2$ et le terme $\langle \nabla f(x), y \rangle$ crée un terme de la forme $-|x|^2$. Bien entendu le résultat ci-dessus va permettre d'établir une bonne partie de nos résultats ou du moins être le point de départ de nos preuves.

Le lemme précédent amène naturellement le premier résultat qui établit la convergence de la suite $(X_n)_{n \geq 0}$ vers un point critique de f .

PROPOSITION 1.3.8. *Si (\mathbf{H}_σ) and (\mathbf{H}_s) sont vérifiées et $(r_n)_{n \geq 1}$ est une suite décroissante vérifiant (\mathbf{H}_r) , alors on a*

(i)

$$\sup_{n \geq 1} \left(\mathbb{E}[f(X_n)] + \frac{1}{r_n} \mathbb{E}[\|Y_n\|^2] \right) < +\infty$$

(ii) $(V_n(X_n, Y_n))_{n \geq 1}$ convergente p.s vers $V_\infty \in \mathbb{R}_+$. En particulier, $(X_n)_{n \geq 1}$ et $(Y_n/\sqrt{r_n})_{n \geq 1}$ sont bornées p.s..

(iii) $\sum_{n \geq 1} \gamma_{n+1} r_n \left(\frac{\|Y_n\|^2}{r_n} + \|\nabla f(X_n)\|^2 \right) < +\infty$ p.s.

(iv) $(Y_n/\sqrt{r_n})_{n \geq 0}$ tends vers 0 quand $n \rightarrow +\infty$ l'ensemble des limites de $(X_n)_{n \geq 0}$ est contenu dans $\{x, \nabla f(x) = 0\}$. De plus, si pour tout z , $\{x, f(x) = z\} \cap \{x, \nabla f(x) = 0\}$ est localement fini, alors $(X_n)_{n \geq 0}$ converge vers un point critique de f .

Une idée de la preuve

Les points (i), (ii) et (iii) se prouvent en appliquant le lemme de Robbins-Siegmund.

Le point (iv) demande lui de recourir à la méthode de l'ODE (voir notamment [20] et la partie sur les algorithmes stochastiques de cette thèse). Bien entendu, les deux mémoires étudiées ne vont pas subir le même traitement. En effet, commençons par étudier le cas de la mémoire exponentielle qui présente la particularité de former un système homogène en temps ; pour rappel on a

$$\begin{cases} X_{n+1} = X_n - \gamma_{n+1} Y_n \\ Y_{n+1} = Y_n + \gamma_{n+1} r (\nabla f(X_n) - Y_n) + \gamma_{n+1} r \Delta M_{n+1}, \end{cases} \quad (1.3.26)$$

Ainsi on peut considérer le système suivant où l'on pose $Z_n = (X_n, Y_n)$,

$$Z_{n+1} = Z_n + \gamma_{n+1} h(Z_n) + \gamma_{n+1} \epsilon_{n+1} \quad (1.3.27)$$

avec

$$h(x, y) = (-y, r(\nabla f(x) - y))^T, \quad \epsilon_{n+1} = (0, r\Delta M_{n+1})^T$$

L'étude suit alors le schéma usuel de preuve induit par la méthode de l'ODE (ref intro) : on construit un processus par interpolation affine et on applique le résultat de Kushner et Clark (voir [72]).

Le cas de la mémoire polynomiale suit le même schéma de preuve mais demande une renormalisation. En effet, l'algorithme est donné par

$$\begin{cases} X_{n+1} = X_n - \gamma_{n+1} Y_n \\ Y_{n+1} = Y_n + \frac{r\gamma_{n+1}}{\Gamma_n} (\nabla f(X_n) - Y_n) + \frac{r\gamma_{n+1}}{\Gamma_n} \Delta M_{n+1}, \end{cases} \quad (1.3.28)$$

Ainsi le principal changement avec le cas exponentiel est l'échelle de temps qu'il faut modifier. En effet, on considère

$$\tilde{\gamma}_n = \gamma_n \sqrt{r_n}, \quad \tilde{\Gamma}_n = \sum_{k=1}^n \tilde{\gamma}_k, \quad \tilde{Y}_n = \frac{Y_n}{\sqrt{r_n}}$$

et on obtient ainsi la procédure suivante qui est homogène en temps

$$\begin{cases} X_{n+1} = X_n - \tilde{\gamma}_{n+1} \tilde{Y}_n \\ \tilde{Y}_{n+1} = \tilde{Y}_n \sqrt{\frac{r_n}{r_{n+1}}} + \tilde{\gamma}_{n+1} \left(\sqrt{\frac{r_n}{r_{n+1}}} \nabla f(X_n) - \frac{r_{n1}}{\sqrt{r_{n+1}}} \tilde{Y}_n + \sqrt{\frac{r_{n1}}{r_{n+1}}} \Delta M_{n+1} \right), \end{cases} \quad (1.3.29)$$

La preuve suit alors le même schéma que le cas exponentiel.

Convergence vers un minimum local

Ici le résultat que nous allons établir va plus loin car si on sait que l'algorithme converge vers un point critique de f , nous n'avons pour le moment aucune garantie de convergence vers un minimum local. Pour résoudre ce problème, il faut fournir à l'algorithme un moyen de contourner les mauvais points critiques appelés *pièges* et c'est le rôle de l'hypothèse (\mathbf{H}_ξ) . Ce genre de problème a été traité par [106] et [20] mais notre algorithme présente la particularité de ne pas avoir de bruit sur sa première composante. Il faut alors adapter les preuves existantes mais l'idée d'origine reste la même : il faut exciter l'algorithme dans les directions répulsives afin qu'il puisse éviter les pièges. Ceci donne lieu au résultat suivant :

THEOREME 1.3.14. *SI f vérifie (\mathbf{H}_s) , le bruit est **elliptique** : i.e. (\mathbf{H}_ξ) est vraie, et la suite $(\gamma_n)_{n \geq 1}$ est telle que (\mathbf{H}_β) et (\mathbf{H}_r) sont satisfaites. Si pour tout z , $\{x, f(x) = z\} \cap \{x, \nabla f(x) = 0\}$ est localement fini, alors :*

- a) *dans le cas $r_n = r$ (mémoire exponentielle), si le bruit a des moments bornés (\mathbf{H}_σ) , alors X_n converge p.s vers un minimum local f .*
- b) *dans le cas $r_n = r\Gamma_n^{-1}$ (mémoire polynomiale), si le bruit est sous-gaussien : i.e. $(\mathbf{H}_{\text{Gauss}, \sigma})$ est vraie, alors X_n converge p.s vers un minimum local f pour tout $\beta < 1/3$.*

Remarque On observe une restriction sur le choix des pas dans le cas de la mémoire polynomiale car l'algorithme est fortement inhomogène en temps en plus de n'avoir un bruit que sur la seconde composante. De plus, il est à noter que dans la plupart des références existantes travaillent avec un bruit borné alors qu'ici il est supposé sous-gaussien ce qui implique la restriction sur β .

Une idée de la preuve

Encore une fois, il va falloir séparer la preuve en fonction du type de mémoire étudiée. Commençons par le cas de la mémoire exponentielle.

Mémoire exponentielle

L'idée est de se ramener au théorème 1 de [33]. Pour cela on reprend la forme vectorielle (1.3.27) et on effectue une linéarisation de la fonction h au voisinage d'un point critique de f que l'on note x_0 . Ainsi, on linéarise la dérive autour de $(x_0, 0) \in \mathbb{R}^d \times \mathbb{R}^d$ comme suit :

$$h(x, y) = \begin{pmatrix} 0 & -I_d \\ rD^2(f)(x_0) & -rI_d \end{pmatrix} \begin{pmatrix} x - x_0 \\ y \end{pmatrix} + O(\|x - x_0\|^2),$$

où I_d la matrice identité et $D^2(f)(x_0)$ la hessienne de f au point x_0 . On diagonalise ensuite la matrice associée à $D^2f(x_0)$ ce qui nous amène à considérer la matrice suivante

$$A_{\lambda, r} := \begin{pmatrix} 0 & -1 \\ r\lambda & -r \end{pmatrix}$$

où λ est une valeur propre de $D^2f(x_0)$ de partie réelle strictement négative. Il suffit pour conclure de remarquer qu'une projection dans une direction répulsive rend le théorème 1 de [33] applicable. C'est ce que permet l'hypothèse H_ξ .

Mémoire polynomiale

Le principe reste le même mais il faut adapter la preuve en tenant compte d'une particularité importante. En effet, le processus est inhomogène en temps, il va donc falloir le renormaliser correctement. Mais il y a aussi des termes de reste qui ne sont pas à une échelle de temps particulière et nécessitent donc un traitement spécifique.

Commençons par effectuer la renormalisation pour la variable Y_n .

$$\tilde{Y}_n = \sqrt{\Gamma_n} Y_n.$$

Ainsi en posant $\tilde{\gamma}_n := \frac{\gamma_{n+1}}{\sqrt{\Gamma_n}}$ on obtient le système suivant

$$\begin{cases} \tilde{X}_{n+1} = \tilde{X}_n - \tilde{\gamma}_{n+1} \tilde{Y}_n \\ \tilde{Y}_{n+1} = \tilde{Y}_n + r\tilde{\gamma}_{n+1} \nabla f(\tilde{X}_n) + \tilde{\gamma}_{n+1} q_{n+1} \Delta M_{n+1} + \tilde{\gamma}_{n+1} U_{n+1} \end{cases} \quad (1.3.30)$$

où $q_{n+1} = \sqrt{\Gamma_{n+1}/\Gamma_n} = 1 + o(n^{-1})$ quand n tend vers l'infini et $(U_n)_{n \geq 1}$ est définie par

$$U_{n+1} = \frac{1/2 - r q_{n+1} + o(n^{-1})}{\sqrt{\Gamma_n}} \tilde{Y}_n + r(q_{n+1} - 1) \nabla f(\tilde{X}_n).$$

Ainsi le système que l'on étudie s'apparente à celui de l'équation différentielle suivante

$$\dot{z}_t = F(z_t)$$

avec

$$F(z) = F(x, y) = (-y, \nabla f(x)).$$

Malheureusement, le théorème 1 de [33] est inapplicable, la faute aux termes qui composent $\tilde{\epsilon}_{n+1}$ car ils sont *trop gros*. Il nous faut alors aller chercher de nouveaux outils.

Afin de gérer les points instables, nous utilisons une fonction de Lyapounov η tirée de [20] qui va permettre de montrer que les point critiques sont des minima locaux. Cependant, l'ingrédient central de la preuve est le lemme de variété stable/instable de Poincaré (voir [108]) sur lequel repose la construction de la fonction η . Cette fonction est de classe C^2 sur $\mathbb{R}^d \times \mathbb{R}^d$, strictement positive et vérifie les propriétés suivantes :

- (i) $\forall z = (x, y) \in \mathcal{N}$, un voisinage compact d'un maximum local de f , $D\eta(z) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ est lipschitzienne, convexe et positivement homogène.
- (ii) Il existe 2 constantes $k, c_1 > 0$ et un voisinage U du point $(0, 0)$ tels que

$$\forall z \in \mathcal{N} \quad \forall u \in U \quad \eta(z + u) \geq \eta(z) + \langle D\eta(z), u \rangle - k\|u\|^2,$$

de plus si $\lfloor \cdot \rfloor_+$ dénote la partie positive alors,

$$\forall z \in \mathcal{N} \quad \forall u \in U \quad \lfloor D\eta(z)(u) \rfloor_+ \geq c_1 \|\Pi_+(u)\|.$$

- (iii) Il existe $\kappa > 0$ tel que

$$\forall z \in \mathcal{N} \quad \langle D\eta(z), F(z) \rangle \geq \kappa \eta(z)$$

Outre ces propriétés qui en font une fonction de Lyapounov, il faut aussi noter que $z \mapsto D\eta(z)$ est une fonction lipschitzienne. Expliciter la fonction est difficile mais on peut néanmoins en dimension 1 voir qu'autour d'un maximum local, on a $\eta(x, y) \simeq \|y - x\|^2$.

On considère ensuite un voisinage \mathcal{N} d'un maximum local de f et on définit le temps d'arrêt suivant

$$T_{\mathcal{N}} = \inf_{n \geq 0} \{ \tilde{Z}_n \notin \mathcal{N} \}$$

et on suppose que $T_{\mathcal{N}} = \infty$. L'objectif est alors d'arriver à une contradiction à l'aide des suites suivantes :

$$\Omega_{n+1} = \eta(\tilde{Z}_{n+1}) - \eta(\tilde{Z}_n), \quad S_n = \eta(\tilde{Z}_n)$$

On montre alors en utilisant les propriétés de la fonction η que S_n ne converge pas vers 0 avec probabilité 1 ce qui contredit le fait que \tilde{Z}_n puisse converger vers un maximum local.

Vitesse de convergence pour les fonctions fortement convexes

Dans le cas où f est fortement convexe, notre étude peut aller plus loin puisque nous sommes capables de donner une borne supérieure pour la vitesse de convergence. C'est ce que résume le résultat suivant.

THEOREME 1.3.15. *Soit x^* l'unique minimum de f et supposons les hypothèses (\mathbf{H}_r) , (\mathbf{H}_s) and $(\mathbf{H}_{sc}(\alpha))$ satisfaites, alors*

- a) *Si $r_n = r$ (mémoire exponentielle) et $\beta < 1$, on a*

$$\mathbb{E}|X_n - x^*|^2 + |Y_n|^2 \lesssim \gamma_n$$

Si $r_n = r$ (mémoire exponentielle) et $\beta = 1$, on a

$$\mathbb{E}|X_n - x^*|^2 + \|Y_n\|^2 \lesssim n^{-(1 \wedge \alpha_r)} \log(n)^{1_{\gamma_{\alpha_r}=1}},$$

où α_r est une constante qui dépend de α et r .

b) Si $r_n = r\Gamma_n^{-1}$ (mémoire polynomiale) et $\beta < 1$, on a

$$\mathbb{E}|X_n - x^*|^2 + \Gamma_n |Y_n|^2 \lesssim \gamma_n$$

Si $r_n = r\Gamma_n^{-1}$ (mémoire exponentielle) et $\beta = 1$, on a

$$\mathbb{E}|X_n - x^*|^2 + n|Y_n|^2 \lesssim \frac{1}{\log n}.$$

Ici le symbole \lesssim signifie que la borne est valable à une constante près (calculable et calculée dans l'article). La distinction des cas $\beta = 1$ et $\beta \in (0, 1)$ provient du terme Γ_n qui vaut respectivement $\log(n)$ et $n^{1-\beta}$. Il est à noter que la vitesse est particulièrement intéressante dans le cas $\beta = 1$ avec la mémoire exponentielle puisque l'on sait depuis les travaux de Nemirovski et Yudin (voir [95]) que la vitesse optimale pour des fonctions fortement convexes est de l'ordre de $1/n$. Cette vitesse est atteignable dans notre cas si $\gamma_{\alpha_r} \geq 1$. On voit ainsi que plus α_r est grand plus la borne est meilleure. Un rapide calcul montre que le choix optimale est $r = 4\alpha$. On a donc tout intérêt à prendre r aussi grand que possible. Il est intéressant de noter que si on prend $r \rightarrow \infty$, on est alors dans un cas proche du gradient stochastique. On peut ainsi constater que si l'on se place dans le cas $\gamma_n = \gamma n^{-1}$ alors une bonne calibration de r ($r \geq 2\alpha$) montre que l'algorithme de boule pesante stochastique a une meilleure vitesse que celle du gradient stochastique. On notera cependant qu'il semble impossible d'obtenir la vitesse $O(1/n)$ indépendamment du choix de (γ, r) , ce que permet la méthode de moyennisation de Ruppert et Polyak.

Une idée de la preuve

Deux schémas de preuve sont proposés dans l'article. Le premier est une étude du cas quadratique où

$$f(x) = \frac{1}{2} \|Ax\|^2$$

avec A une matrice symétrique. Cette étude étant très calculatoire nous ne la détaillons pas ici mais le principe est simple : il s'agit de travailler avec une écriture vectorielle du système en posant $Z_n = (X_n, Y_n)$,

$$Z_{n+1} = Z_n + \gamma_{n+1} A_n(Z_n) + \gamma_{n+1} \epsilon_{n+1} \quad (1.3.31)$$

où

$$A_n(x, y) = \begin{pmatrix} 0 & -1 \\ r_n & -r_n \end{pmatrix}$$

et

$$\epsilon_{n+1} = \begin{pmatrix} 0 \\ \Delta M_{n+1} \end{pmatrix}$$

On montre alors que les valeurs propres cette matrice sont de partie réelles strictement négatives. Bien entendu, il faudra distinguer les cas exponentiels et polynomiaux qui diffèrent fortement.

Le cas non quadratique se gère à l'aide de la fonction V_n introduite précédemment. L'idée centrale

est d'utiliser la formule suivante (tirée du Lemme 1.3.2)

$$\mathbb{E}(V_{n+1}(X_{n+1}, Y_{n+1}) | \mathcal{F}_n) \leq V_n(X_n, Y_n) + \gamma_{n+1} \Delta_{n+1}$$

où

$$\Delta_{n+1} = -c_{a,b}|Y_n|^2 - br_n|\nabla f(X_n)|^2 + \Delta r_{n+1}$$

L'idée est alors d'utiliser une montée en puissance pour le terme

$$u_{n+1} := (V_n(X_n, Y_n) + \gamma_{n+1} \Delta_{n+1})^K$$

L'objectif est alors d'obtenir une récurrence de la forme

$$u_{n+1} \leq u_n(1 - C_K r_n + C r_n \gamma_{n+1}^2) + C r_n \gamma_{n+1}^2$$

où C_K est une constante dépendante de K .

Ainsi, il nous reste à calibrer K afin d'obtenir le résultat. L'étude se résume comme suit. Considérons $\gamma_n = \gamma n^{-\beta}$ avec $\beta \in (0, 1]$

- Si $r_n = r$ et $\beta < 1$, il suffit de prendre $K = 1$.
- Si $r_n = r$ et $\beta = 1$, il faut choisir $C_K \gamma > 1$.
- Si $r_n = r/\Gamma_n$ et $\beta < 1$, il faut alors choisir K suffisamment grand pour garantir l'existence d'un réel $\rho > 1$ et d'un rang $n_0 \geq 1$ tels que

$$\forall n \geq n_0, \quad u_{n+1} \leq (1 - \frac{\rho}{n})u_n + C n^{-\beta-1}.$$

On a alors

$$u_n \leq u_{n_0} \prod_{k=n_0}^n (1 - \frac{\rho}{k}) + C \sum_{k=n_0+1}^n k^{-\beta-1} \prod_{\ell=k+1}^n (1 - \frac{\rho}{\ell}).$$

En utilisant $1 - x \leq \exp(-x)$ on obtient finalement

$$u_n \leq C n^{-\rho} (1 + \sum_{k=n_0+1}^n k^{-\beta-1+\rho}) \leq C n^{-\beta}$$

Un résultat de convergence en loi

Dans le cas de la mémoire exponentielle, nous parvenons à obtenir un théorème central limite. Nous supposons que f a un unique minimum en $x^* = 0$.

THEOREME 1.3.16. *Sous les hypothèses du théorème (1.3.15), si $r_n = r$ et $\beta < 1$ ou si $\beta = 1$ et $\gamma \alpha_r > 1$, supposons que*

$$\mathbb{E}[(\Delta M_{n+1})(\Delta M_{n+1})^t | \mathcal{F}_{n-1}] = \mathbb{E}[(\Delta M_{n+1})(\Delta M_{n+1})^t | X_n = x] \xrightarrow{n \rightarrow +\infty} \sigma^2 \quad \text{en probabilité} \quad (1.3.32)$$

où σ^2 est une matrice symétrique définie positive $d \times d$. Alors :

- i) l'algorithme renormalisé $\left(\frac{X_n}{\sqrt{\gamma_n}}, \frac{Y_n}{\sqrt{\gamma_n}}\right)_n$ converge vers la mesure invariante $\mu_\infty^{(\beta)}$ de la diffu-

sion dont le g  n  rateur infinit  simal \mathcal{L} est d  fini pour les fonctions de classe \mathcal{C}^2 par :

$$\mathcal{L}g(z) = \left\langle \nabla g(z), \left(\frac{1}{2\gamma} 1_{\{\beta=1\}} I_{2d} + H \right) z \right\rangle + \frac{1}{2} \text{Tr}(\Sigma^T D^2 g(z) \Sigma)$$

avec

$$H = \begin{pmatrix} 0 & -I_d \\ rD^2 f(x^*) & -rI_d \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \sigma(x^*) \end{pmatrix}.$$

ii) Si $\sigma^2(x) = \sigma^2 I_d$ et $\beta < 1$, alors $\mu_\infty^{(\beta)}$ est une mesure gaussienne centr  e et de matrice de covariance

$$\frac{\sigma^2}{2} \begin{pmatrix} \{D^2 f(x^*)\}^{-1} & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & rI_d \end{pmatrix}$$

En particulier,

$$\frac{X_n}{\sqrt{\gamma_n}} \Longrightarrow \mathcal{N}\left(0, \frac{\sigma^2}{2} \{D^2 f(x^*)\}^{-1}\right).$$

iii) Si $\gamma_{\alpha_r} > 1$ et $\sigma^2(x) = \sigma^2 I_d$ et $\beta = 1$, $\mu_\infty^{(\beta)}$ est une mesure gaussienne centr  e et de matrice de covariance d  pendante de $r, D^2 f(x^*)$ et γ .

Une id  e de la preuve

Nous commen  ons par renormaliser correctement l'algorithme en consid  rant $(\bar{Z}_n)_{n \geq 1}$ d  finie par :

$$\bar{Z}_n = \frac{Z_n}{\sqrt{\gamma_n}} = \left(\frac{X_n}{\sqrt{\gamma_n}}, \frac{Y_n}{\sqrt{\gamma_n}} \right).$$

f   tant de classe \mathcal{C}^2 (et $x^* = 0$), on lin  arise ∇f au voisinage de 0 en utilisant une formule de Taylor    l'ordre 2, on montre l'existence d'un r  el $\xi_n \in [0, X_n]$ tel que

$$\nabla f(X_n) = D^2 f(\xi_n) X_n.$$

On a alors

$$\bar{Z}_{n+1} = \bar{Z}_n + \gamma_{n+1} b_n(\bar{Z}_n) + \sqrt{\gamma_{n+1}} \begin{pmatrix} 0 \\ \Delta M_{n+1} \end{pmatrix}$$

o   b_n est d  fini par

$$b_n(z) = \frac{1}{\gamma_{n+1}} \left(\sqrt{\frac{\gamma_n}{\gamma_{n+1}}} - 1 \right) z + \bar{C}_n z, \quad z \in \mathbb{R}^{2d}, \quad (1.3.33)$$

avec

$$\bar{C}_n := \sqrt{\frac{\gamma_n}{\gamma_{n+1}}} \begin{pmatrix} 0 & -I_d \\ rD^2 f(\xi_n) & -rI_d \end{pmatrix}. \quad (1.3.34)$$

Il est important de noter que

$$\frac{1}{\gamma_{n+1}} \left(\sqrt{\frac{\gamma_n}{\gamma_{n+1}}} - 1 \right) = \gamma^{-1} (n+1)^\beta \left[1 + \frac{\beta}{2n} + o(n^{-1}) - 1 \right] = \begin{cases} o(n^{\beta-1}) & \text{if } \beta < 1 \\ \frac{1}{2\gamma} + o(1) & \text{if } \beta = 1 \end{cases} \quad (1.3.35)$$

On d  fini alors,    partir de la suite $(\bar{Z}_n)_{n \geq 1}$, la suite $(\bar{Z}^{(n)})_{n \geq 1}$ de processus    temps continu

définit par

$$\bar{Z}_t^{(n)} = \bar{Z}_n + B_t^{(n)} + M_t^{(n)}, \quad t \geq 0, \quad (1.3.36)$$

avec

$$B_t^{(n)} = \sum_{k=n+1}^{N(n,t)} \gamma_k b_{k-1}(\bar{Z}_{k-1}) + (t - \underline{t}_n) b_{N(n,t)}(\bar{Z}_{N(n,t)}),$$

$$M_t^{(n)} = \sum_{k=n+1}^{N(n,t)} \sqrt{\gamma_k} \begin{pmatrix} 0 \\ \Delta M_k \end{pmatrix} + \sqrt{t - \underline{t}_n} \begin{pmatrix} 0 \\ \Delta M_{N(n,t)+1} \end{pmatrix}.$$

Rappelons les définitions standards utilisées ci-dessus $\underline{t}_n = \Gamma_{N(n,t)} - \Gamma_n$ où $N(n,t) = \min \left\{ m \geq n, \sum_{k=n+1}^m \gamma_k > t \right\}$.

Pour obtenir le résultat, nous montrerons que $(\bar{Z}^{(n)})_{n \geq 1}$ converge vers une diffusion stationnaire, nous procéderons en deux étapes en montrant la tension de la suite (une forme de compacité probabiliste) puis nous identifierons la limite en utilisant un problème de martingale.

1.4 Approximation du régime stationnaire du processus de McKean-Vlasov

L'approximation de la mesure invariante d'un processus est une question dont l'intérêt pratique trouve sa source dans des domaines divers comme la mécanique. Les méthodes permettant de telles études reposent sur les propriétés de convergence vers l'état d'équilibre. Principalement, on peut penser à deux méthodes permettant d'illustrer cette convergence. Tout d'abord, on peut utiliser la convergence en loi du semi-groupe associé. Dans ce cadre, l'idée est de fixer t grand puis d'approcher $P_t(x, \cdot)$ via une méthode de Monte-Carlo impliquant la simulation d'un grand nombre de trajectoires. La seconde approche naturelle consiste à approximer la mesure d'occupation ("temporelle") du processus et est alors basée sur la simulation d'une seule trajectoire. C'est cette seconde approche que nous privilégierons ici pour l'approximation de la mesure invariante de l'EDS de type McKean-Vlasov.

Dans le cadre des diffusions, Talay [118] propose un algorithme basé sur un schéma d'Euler et de Milstein à pas constant. Cette technique fait intervenir deux erreurs : l'erreur de discrétisation et celle induite par le comportement en temps long. Afin de gérer conjointement ces erreurs, les travaux de Lamberton et Pagès [77] proposent d'introduire un pas décroissant et ainsi de se ramener à l'étude d'un algorithme stochastique. Ces résultats ont ensuite été généralisés par Lemaire [82] et Panloup [104] pour des processus à dérive non-lipschitzienne et à des équations différentielles stochastiques dirigées par des processus de Levy. Nos travaux s'inscrivent dans un cadre différent puisque nous allons tenter d'approcher la mesure invariante d'un processus non markovien.

Le processus de McKean-Vlasov est un processus dont le comportement est décrit par une équation différentielle dont les coefficients dépendent de la loi de la solution. Étudié pour la première fois par Henry McKean en 1966 (voir [89]), ce processus peut être vu comme un modèle simplifié de l'équation de Boltzmann qui décrit de la densité de particules en interaction. Ce processus a attiré l'attention de nombreux travaux du fait de ses particularités structurelles. On trouvera entre autre des travaux sur l'existence de mesure invariante ou encore les vitesses de convergence à l'équilibre. Sa structure particulière rend ces études difficiles mais la propagation du chaos permet de ramener l'étude de l'équation différentielle stochastique à celle d'un système de particules en interactions. L'heuristique derrière la propagation du chaos se niche dans l'asymptotique du système considéré. En effet, on considère N particules en interactions les unes avec les autres et en faisant tendre N vers l'infini ce système se comporte comme l'équation différentielle de McKean-Vlasov.

Ces considérations vont nous amener à étudier un algorithme stochastique associé à l'équation de McKean-Vlasov pour lequel nous montrerons qu'il offre un moyen d'approcher la mesure invariante. Comme nous l'avons dit, les méthodes existantes pour les diffusions ne sont pas applicables pour ce processus mais nous verrons que le processus de McKean-Vlasov se rapproche fortement de la classe des processus auto-agissants. Ces processus sont des diffusions définies par une équation différentielle stochastique dont le terme de dérive fait intervenir tout le passé du processus. Nous profiterons de la proximité du processus de McKean-Vlasov avec les diffusions auto-agissants et utiliserons les travaux de [70] et [23] afin de proposer un résultat de convergence faible de la mesure d'occupation d'une version discrétisée de l'EDS. Comme nous le verrons la principale difficulté se trouve dans la caractérisation de la limite. Dans la lignée des travaux précédemment cités, l'idée sera alors d'identifier une dynamique limite de type EDO sur l'espace des probabilités. On fera alors appel à la notion de pseudo-trajectoire asymptotique, concept dont on introduira la définition dans la suite.

Contenu de la sous-partie

Dans cette partie, notre objectif est de présenter les résultats de l'article en cours de rédaction *Computation of the invariant distribution for stochastic McKean-Vlasov equation*. Dans cette optique, nous présenterons les équations d'intérêt en donnant les principales propriétés connues. Nous proposerons l'algorithme stochastique à partir duquel nous construirons la mesure d'occupation dont nous étudierons la convergence. La particularité de notre étude par rapport à celles déjà existantes est la non-linéarité du processus de McKean-Vlasov qui nous poussera à utiliser des méthodes développées pour l'étude des processus auto-agissants. Nous présenterons ainsi la notion de pseudo-trajectoire asymptotique qui permet de caractériser la limite de notre algorithme stochastique.

1.4.1 Approximation du régime stationnaire de diffusions

C'est au début des années 90 que les premiers travaux dans le domaine sont apparus sous l'impulsion de Talay notamment qui proposa une approximation basée sur le contrôle de deux erreurs. Tout d'abord, il considère un schéma d'Euler à pas constant γ associé à la diffusion. Ensuite, une suite de mesures d'occupations $(\nu_n^\gamma)_{n \geq 1}$ associées au schéma d'Euler précédemment défini. Sous de bonnes hypothèses, la convergence vers la mesure invariante est obtenue en faisant tendre n vers l'infini et γ vers 0.

En introduisant un schéma d'Euler à pas décroissant, Lamberton et Pagès proposèrent une étude basée sur un algorithme stochastique. L'idée est de considérer une suite de pas de temps $(\nu_n^\gamma)_{n \geq 1}$ décroissante et qui tend vers 0 quand n tend vers l'infini. En plus de montrer la convergence vers la mesure invariante, les auteurs obtiennent des vitesses de convergences et un théorème limite central en utilisant pour l'essentiel des méthodes de martingales. Rappelons maintenant quelques uns de ces résultats ainsi que les hypothèses standards.

Considérons un processus $(X_t)_{t \geq 0}$ à valeurs dans \mathbb{R}^d solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dt \quad (1.4.37)$$

et sa mesure invariante ν supposée unique. On suppose également que les coefficients b et σ sont continus et localement lipschitziens i.e

$$\forall x, y \in \mathbb{R}^d \quad \text{tels que} \quad |x| \leq c \quad |y| \leq c, \quad |b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq K_c |x - y|$$

Sous ses hypothèses, le processus $(X_t)_{t \geq 0}$ est markovien de semi-groupe $(P_t)_{t \geq 0}$ et de générateur L définit pour toute fonction $f \in C^2(\mathbb{R}^d)$ par

$$Lf(x) = \frac{1}{2} \text{Tr}(\sigma(x)D^2 f(x)\sigma^*(x)) + \langle b(x), \nabla f(x) \rangle, \quad x \in \mathbb{R}^d$$

L'étude porte alors sur un algorithme stochastique associé à (1.4.37) que l'on définit comme suit. Tout d'abord, on considère une suite de pas décroissante $(\gamma_n)_{n \geq 0}$ telle que :

$$\lim_{n \rightarrow \infty} \gamma_n = 0$$

On définit aussi la suite des temps de discrétisation $(\Gamma_n)_{n \geq 0}$ par

$$\Gamma_n = \sum_{k=1}^n \gamma_k.$$

On définit alors un schéma d'Euler $(\bar{X}_t)_{t \geq 0}$ pour la diffusion $(X_t)_{t \geq 0}$ comme suit.

$$\begin{aligned} \bar{X}_0 &= x, \\ \bar{X}_{\Gamma_{n+1}} &= \bar{X}_{\Gamma_n} + \gamma_{n+1} b(\bar{X}_{\Gamma_n}) + \sqrt{\gamma_{n+1}} U_{n+1}, \quad n \geq 1 \end{aligned}$$

où $(U_n)_{n \geq 1}$ est un bruit blanc.

On associe à ce schéma une suite de mesure empirique $(\bar{\nu}_n)_{n \geq 1}$ définie par

$$\bar{\nu}_n(dx) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{\bar{X}_{\Gamma_{k-1}}}(dx)$$

où $(\eta_n)_{n \geq 1}$ est une suite de réels strictement positifs avec $H_n = \sum_{k=1}^n \eta_k$.

Remarque Une question naturelle ici est celle du choix des suites $(\gamma_n)_{n \geq 1}$ et $(\eta_n)_{n \geq 1}$. Nous nous placerons dans le cas pratique où $\gamma_n = \eta_n$ qui s'interprète comme la discrétisation naturelle de la mesure d'occupation de $(X_t)_{t \geq 0}$. La question sur d'autre choix de pas est étudiée notamment par Lemaire [82] et l'on renvoie le lecteur à ces travaux pour de plus amples informations.

Nous commençons par donner une idée des résultats qui peuvent être obtenus en nous appuyant sur [77]. Nous présentons dans un premier temps un résultat de convergence sous l'hypothèse de l'existence d'une fonction de Lyapounov puis un résultat en lien avec la vitesse de convergence de telles procédures. Pour une matrice symétrique $A \in \mathcal{M}_d$, on notera $\lambda_A^+ = \max(\lambda_1, \dots, \lambda_d)$ où $\lambda_1, \dots, \lambda_d$ sont les valeurs propres de A .

Les hypothèses suivantes vont permettre de donner des résultats de convergence.

(H1) Il existe une constante C_a positive telle que : $|b|^2 + \text{Tr}(\sigma\sigma^*) \leq C_a V^a$.

(H2) Il existe des constantes $c, d > 0$ tels que : $\langle \nabla V, b \rangle + \lambda_p \text{Tr}(\sigma\sigma^*) \leq c - dV^a$.

où $\lambda_p = \frac{1}{2} \sup_{x \in \mathbb{R}^d} \lambda^+$
 $D^2V + (p-1) \frac{\nabla V \otimes \nabla V}{V}$. Alors,

PROPOSITION 1.4.9. *Supposons les hypothèses (H1) et (H2) satisfaites pour $a \in (0, 1]$ et $p > 2$. Alors,*

— on a

$$\sup_{n \geq 1} \bar{\nu}_n(w, V^{\frac{p}{2}+a-1}) < \infty$$

En particulier, ceci implique la tension de la suite $(\bar{\nu}_n(w, dx))_{n \geq 1}$.

— *Toute valeur d'adhérence de $(\bar{\nu}_n(w, dx))_{n \geq 1}$ est p.s une probabilité invariante pour $(X_t)_{t \geq 1}$. En particulier, lorsque cette mesure est unique alors pour toute fonction $f \in C(\mathbb{R}^d)$ telle que $f(x) \simeq V^{p/2+a-1}(x)$ on a*

$$\bar{\nu}_n(w, f) \rightarrow_{n \rightarrow \infty} \nu(f).$$

Ce résultat garantit la convergence pour une certaine classe de fonctions mais ne donne aucune idée de la vitesse. Pour quantifier cela, on s'appuie sur l'équation de Poisson associé à L définit comme suit

$$f - \nu(f) = Lg \quad (1.4.38)$$

Sous des hypothèses d'existence d'une solution notée g_f de (1.4.38), on est alors capable de quantifier la vitesse.

Nous noterons (\mathbf{C}_f^5) l'hypothèse suivante :

- Il existe une solution g_f de classe $C^5(\mathbb{R}^d)$ telle que les dérivées partielles d'ordre inférieure ou égale à 5 soit dominées par V^5 .

Pour des résultats sur l'existence de solutions pour l'équation de Poisson et leurs régularités nous renvoyons le lecteur à [74].

Pour comprendre le besoin d'une telle fonction, appliquons la formule d'Itô à la fonction g . On obtient ainsi

$$g(X_t) = g(X_0) + \int_0^t Lg(X_s)ds + \int_0^t \langle \nabla g(X_s), \sigma(X_s) \rangle dB_s \quad (1.4.39)$$

Comme g est solution de (1.4.38) on peut réécrire cela comme suit

$$\frac{1}{t} \int_0^t f(X_s)ds - \nu(f) = \frac{1}{t} \int_0^t Lg(X_s)ds = \frac{g(X_t) - g(X_0)}{t} - \underbrace{\frac{1}{t} \int_0^t \langle \nabla g(X_s), \sigma(X_s) \rangle dB_s}_{\text{martingale brownienne}}$$

Ainsi, en gérant de manière adéquate le premier terme du membre de droite et en utilisant un TCL pour la martingale brownienne, on peut obtenir le résultat suivant

$$\frac{1}{\sqrt{t}} \left(\int_0^t f(X_s)ds - \nu(f) \right) \Rightarrow \mathcal{N}(0, \sigma_f^2)$$

Le théorème suivant tiré et adapté de [77] illustre cette idée.

THEOREME 1.4.17. *Supposons que*

$$|b|^2 \leq C_a V^a, \quad \{\nabla V, b\} \leq \kappa_1 - \kappa_2 V^a, \quad Tr(\sigma\sigma^*) = o(V^a)$$

avec $a \in (0, 1]$, $\kappa_1, \kappa_2 > 0$. *Supposons de plus qu'il existe $f : \mathbb{R}^d \mapsto \mathbb{R}$ vérifiant (\mathbf{C}_f^5) , alors*

$$\text{— Si } \frac{\sum_{k=1}^n \gamma_k^2}{\sqrt{\Gamma_n}} \rightarrow_{n \rightarrow \infty} 0,$$

$$\sqrt{\Gamma_n}(\bar{\nu}_n(w, f) - \nu(f)) \Rightarrow \mathcal{N}(0, \int_{\mathbb{R}^d} |\sigma^* \nabla g_f|^2 d\nu) \quad \text{lorsque } n \rightarrow \infty$$

On peut aussi obtenir un résultat dans le cas où $\frac{\sum_{k=1}^n \gamma_k^2}{\sqrt{\Gamma_n}}$ converge vers un réel $\beta > 0$, nous renvoyons le lecteur à [77] pour un résultat en ce sens. Ce résultat permet d'obtenir une vitesse en explicitant le terme $\sqrt{\Gamma_n}$. Pour cela considérons un pas de la forme $\gamma_n \simeq n^{-\theta}$ ce qui entraîne alors $\Gamma_n \simeq n^{1-\theta}$ si $\theta \neq 1$ ou alors $\Gamma_n \simeq \log(n)$ si $\theta = 1$. On note alors que

$$\frac{\sum_{k=1}^n \gamma_k^2}{\sqrt{\Gamma_n}} \rightarrow_{n \rightarrow \infty} 0 \Leftrightarrow \theta > 1/3$$

ceci entraîne une vitesse de convergence maximale pour le choix $\theta = 1/3$ avec une vitesse de l'ordre de $n^{1/3}$.

Remarque Des développements sur ce thème ont été proposés par Pagès et Panloup (voir [104], [100] et [101]) avec une généralisation "fonctionnelle" de la suite de mesure d'occupation afin d'approcher le régime stationnaire global. Le lecteur trouvera dans [104], un algorithme basé sur une mesure d'occupation moins moyennisantes afin de réduire le temps de calcul.

1.4.2 Le processus de McKean-Vlasov

Les équations de McKean-Vlasov sont à l'origine dues à Anatoly Vlasov qui les utilisa pour modéliser la dynamique de particules de plasma en interaction les unes avec les autres. Plus tard en 1966, Henry McKean s'intéressa à ces équations en montrant notamment le lien avec les équations de Boltzmann. Commençons par rappeler la définition du processus de McKean-Vlasov. On fixe un potentiel d'interaction $W : \mathbb{R}^{2d} \rightarrow \mathbb{R}^+$, un potentiel de confinement $V : \mathbb{R}^d \rightarrow \mathbb{R}^+$ et on considère l'équation différentielle stochastique suivante :

$$dX_t = -\nabla V(X_t)dt - \nabla_x W * \nu_t(X_t)dt + \sigma dB_t \quad (1.4.40)$$

$$\nu_t = \mathcal{L}(X_t) \quad (1.4.41)$$

où $(B_t)_{t \geq 0}$ est un mouvement brownien, avec $X_0 \in \mathbb{R}$.

L'existence de solutions pour cette équation n'est pas une question triviale. En effet, le processus est fortement non-linéaire au sens où la dérive fait intervenir la loi de la solution. De plus, le processus n'est pas markovien (il a de la mémoire) néanmoins le couple $(X_t, \nu_t)_{t \geq 0}$ l'est. On trouvera dans [120] et [55] des résultats d'existence en supposant notamment V et W convexes. Pour faciliter l'étude de ce processus on considère souvent un système de N particules $(X_t^i)_{1 \leq i \leq N}$ qui sont solutions des équations différentielles suivantes

$$dX_t^i = -\nabla V(X_t^i)dt - \frac{1}{N} \left(\sum_{j=1}^N \nabla_x W(X_t^i, X_t^j) \right) dt + \sigma dB_t^i \quad X_0^i = x_0 \in \mathbb{R} \quad (1.4.42)$$

Le lien entre ces équations et le processus de McKean-Vlasov tient dans la notion de propagation du chaos. En effet, si on fait tendre le nombre de particules N vers l'infini, on peut espérer qu'une particule prise au hasard se comporte comme une solution de l'équation de McKean-Vlasov.

Cette définition est à l'origine du travail de Bossy et Talay [29] qui ont proposé un moyen d'approcher le processus de McKean-Vlasov en utilisant notamment la notion de propagation du chaos

comme moyen d'estimer les erreurs d'approximation dues à l'utilisation du schéma particulière (1.4.42).

Avant de présenter notre méthode, il convient de rappeler des résultats d'existence à propos de mesures invariantes pour l'équation de McKean-Vlasov. Il existe des configurations pour lesquelles on peut avoir plusieurs mesures invariantes comme le montre le résultat suivant extrait de [64].

THEOREME 1.4.18. *Si V'' est convexe alors il existe exactement trois mesures invariantes.*

Pour des raisons pratiques, nous travaillerons dans le cadre où la mesure invariante est unique. C'est ce qu'illustre le résultat de [38].

THEOREME 1.4.19. *Si V et W sont des fonctions fortement convexes alors le processus de McKean-Vlasov admet une unique mesure invariante.*

Parmi les autres résultats on peut citer les travaux de [92] et [38] qui s'attachent à trouver des vitesses de convergence pour des processus dits de Vlasov-Fokker-Planck (une version cinétique de l'équation de McKean-Vlasov).

1.4.3 Les diffusions auto-agissantes

Notre étude prend une autre direction en s'appuyant sur les travaux de [23] et [70] où l'étude du comportement est faite pour des processus auto-agissants définis au travers de l'EDS suivante

$$dY_t = \sigma dB_t + \left(\int_0^t f(Y_t, Y_s) ds \right) dt \quad (1.4.43)$$

où $(B_t)_{t \geq 0}$ est un mouvement brownien, avec $X_0 \in \mathbb{R}$ et f une fonction à valeurs réelles. On peut rapprocher les processus de McKean-Vlasov de la classe des processus auto-agissant au sens où il s'agit de processus qui satisfont une équation différentielle faisant intervenir le passé du processus. Pour ce type de processus, on trouve une littérature assez dense. Le lecteur intéressé pour se tourner vers les travaux de [38] et ceux de [92] où une étude portant sur les vitesses de convergence à l'équilibre pour des processus de Vlasov-Fokker-Planck sont données. A l'origine de l'étude du comportement en temps long de tels processus, on trouve les travaux de [42] qui s'intéressent à la convergence presque sûre des diffusions auto-agissantes. Pour un panorama plus complet on se tournera vers le travail de Pemantle [105] et les références qu'il contient.

Les deux travaux qui ont inspiré grandement notre travail sont ceux de Benaim, Ledoux et Raimond [23] d'une part et de Kurtzmann [70] d'autre part. En effet, ces deux articles s'intéressent à l'étude de processus auto-agissants respectivement dans les cas compact et non-compact. Dans [23] est étudié un processus de la forme :

$$dX_t = -\nabla_x W * \nu_t(X_t) dt + \sigma dB_t \quad (1.4.44)$$

avec $W : \mathbb{R}^{2d} \mapsto \mathbb{R}^+$ et

$$\nu_t = \frac{1}{t} \int_0^t \delta_{X_s} ds.$$

où δ_x est la masse de Dirac au point x . Le symbole $*$ désigne la convolée et est défini comme suit

$$f * \nu(x) := \int_{\mathbb{R}^d} f(x, y) \nu(dy)$$

En d'autres termes, on fait interagir le processus avec son passé au travers du potentiel W .

L'étude porte sur le comportement en temps de la suite de mesure d'occupation $(\nu_t)_{t \geq 0}$. Les auteurs montrent que cette dernière converge presque sûrement en utilisant la notion de pseudo-trajectoire asymptotique introduite par [21] pour caractériser la limite. Les travaux de [70] ont étendu ce résultat à \mathbb{R}^d en étudiant le processus suivant

$$dX_t = -\nabla V(X_t)dt - \nabla_x W * \nu_t(X_t)dt + \sigma dB_t \quad (1.4.45)$$

On notera l'apparition du terme V qui vient compenser la perte de compacité par rapport au cadre de [23]. Comme dans le travail de [23], nous aurons besoin de travailler avec une version linéarisée de (1.4.40) en un sens que nous définissons ci-dessous.

L'idée est de considérer une version linéarisée de (1.4.45) en gelant la mesure ν_t . On obtient ainsi l'équation :

$$dX_t = -\nabla V(X_t)dt - \nabla_x W * \nu(X_t)dt + \sigma dB_t \quad (1.4.46)$$

Le système devient alors une diffusion standard que l'on note X^ν pour laquelle on peut définir le semi-groupe associé $(P_t^\nu)_{t \geq 0}$ ainsi que son générateur infinitésimal A_ν . On peut aussi définir l'opérateur Q_ν , solution de l'équation de Poisson associé à A_ν et défini par la relation suivante :

$$A_\nu \circ Q_\nu = \Pi_\nu - Id$$

où Π_ν est la mesure invariante de (1.4.39). Pour illustrer l'intérêt de cet opérateur, on applique la formule d'Îto.

$$Q_\nu f(X_t^\nu) = Q_\nu f(x) + \int_0^t \nabla Q_\nu f(X_s^\nu) dB_s + \int_0^t \nabla A_\nu \circ Q_\nu f(X_s^\nu) ds \quad (1.4.47)$$

Ainsi en utilisant la définition de Q_ν , on obtient

$$\frac{1}{t}(-Q_\nu f(X_t^\nu) + Q_\nu f(x) + \int_0^t \nabla Q_\nu f(X_s^\nu) dB_s) = \frac{1}{t} \int_0^t (\delta_{X_s^\nu} - \Pi_\nu f) ds \quad (1.4.48)$$

En résumé, si on parvient à montrer que le terme de gauche tends vers 0 quand t tends vers l'infini alors on arrive à prouver que ν_t tend vers Π_ν .

Le cas de l'équation (1.4.45) est plus technique car la mesure dépend du temps mais l'idée sous-jacente est celle exposée ci-dessus. Pour parvenir au résultat, il est attendu comme pour la plupart des processus stochastiques que la trajectoire de ν_t mime celle d'un flot déterministe. C'est d'ailleurs cette idée qui guide le travail de [23].

Le concept associé est celui de pseudo-trajectoire que l'on définit ici.

DEFINITION 1.4.6. Une fonction continue $X : \mathbb{R}^+ \mapsto \mathcal{P}(\mathbb{R}^d)$ est une pseudo-trajectoire asymptotique pour le semi-flot φ si pour tout $T > 0$ on a

$$\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq T} d(X_{t+s}, \varphi_s(X_t)) = 0$$

où d est définie en considérant une suite de fonctions $(f_k)_{k \geq 0}$ de $C^\infty(\mathbb{R}^d)$ denses dans l'espace des fonctions continues. On a alors

$$d(\nu, \mu) := \sum_{k \geq 0} \frac{1}{2^k} |\nu(f_k) - \mu(f_k)|.$$

La définition d'une pseudo-trajectoire asymptotique signifie qu'asymptotiquement en temps, le semi-flot φ coïncide avec la trajectoire shiftée du processus X .

Les travaux de Kurtzmann ont notamment nécessité un moyen de recréer de la compacité afin de pouvoir travailler dans un cadre plus aisé structurellement. Le potentiel V est évidemment la pièce maitresse de ce travail et l'idée est d'introduire la norme d'opérateur appelé V -norme qui est définie par

$$\|f\|_V = \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{V(x)}$$

On lui associe l'espace des fonctions continues et V -bornées

$$C^0(\mathbb{R}^d; V) = \{f \in C^0(\mathbb{R}^d); \|f\|_V < \infty\}$$

De manière similaire, nous avons

$$C^\infty(\mathbb{R}^d; V) = C^\infty(\mathbb{R}^d) \cap C^0(\mathbb{R}^d; V).$$

Nous définissons un sous-espace des mesures de Borel signées sur \mathbb{R}^d ,

$$\mathcal{M}(\mathbb{R}^d; V) := \{\nu \in \mathcal{M}(\mathbb{R}^d); \int_{\mathbb{R}^d} V(y) |\nu|(dy) < \infty\}$$

où $|\nu|$ est la variation de ν : $|\nu| := \nu^+ + \nu^-$ dont la définition provient de la décomposition de Hahn-Jordan. La norme

$$\|\nu\|_V := \sup_{f \in C^0(\mathbb{R}^d; V); \|f\|_V \leq 1} \left| \int_{\mathbb{R}^d} f d\nu \right|,$$

fait de notre espace de probabilité un espace de Banach. Nous nous restreindrons à l'espace des mesures de probabilités suivant $\mathcal{P}(\mathbb{R}^d; V) = \mathcal{M}(\mathbb{R}^d; V) \cap \mathcal{P}(\mathbb{R}^d)$. La topologie forte de $\mathcal{P}(\mathbb{R}^d; V)$ est la topologie trace de $\mathcal{M}(\mathbb{R}^d; V)$. Elle fait de $\mathcal{P}(\mathbb{R}^d; V)$ un espace métrique complet. En particulier, la convergence faible est définie pour une classe de fonctions plus large que celle de la définition standard :

$$\nu_n \rightarrow \nu, \quad \text{si et seulement si} \quad \int_{\mathbb{R}^d} f d\nu_n \rightarrow_{n \rightarrow \infty} \int_{\mathbb{R}^d} f d\nu \quad \forall f \in C^0(\mathbb{R}^d; V).$$

Enfin, pour tout $\beta > 1$, on introduit l'espace suivant

$$\mathcal{P}_\beta(\mathbb{R}^d; V) := \{\nu \in \mathcal{P}(\mathbb{R}^d; V); \int_{\mathbb{R}^d} V(y)\nu(dy) \leq \beta\}$$

dont l'intérêt se trouve dans la proposition suivante extraite de [70].

PROPOSITION 1.4.10. *Soit $\beta > 1$. L'ensemble $\mathcal{P}_\beta(\mathbb{R}^d; V)$ est un sous-ensemble faiblement compact de $\mathcal{P}(\mathbb{R}^d; V)$.*

On a donc ainsi un moyen de compactifier l'espace des probabilités (dans lequel vit asymptotiquement $(\nu)_{t \geq 0}$).

1.4.4 Approximation de la mesure invariante pour l'équation de McKean-Vlasov

L'objectif de cette sous-section est de présenter nos résultats. Nous commencerons par présenter l'algorithme que nous étudierons avec une présentation des diverses hypothèses entrant en jeu. Nous établirons notre résultat au moyen de deux résultats préliminaires. Un premier établira la tension de notre processus et le second établira la caractérisation de la limite de ce dernier au moyen de la notion de pseudo-trajectoire asymptotique.

Définitions et premiers résultats

Commençons par présenter l'algorithme que nous allons étudier. On définit $(X_n)_{n \geq 0}$ le schéma d'Euler de pas $(\gamma_n)_{n \geq 0}$ associé à la diffusion $(X_t)_{t \geq 0}$ solution de (1.4.40) comme suit,

$$X_{n+1} = X_n - \gamma_{n+1} \left(\nabla V(X_n) + \frac{1}{\Gamma_n} \sum_{k=0}^n \gamma_{k+1} \nabla_x W(X_n, X_k) \right) + \sqrt{\gamma_{n+1}} U_{n+1} \quad (1.4.49)$$

où $X_0 \in \mathbb{R}$ et $(U_n)_{n \geq 1}$ est un bruit blanc possédant un moment quatrième.

On définit ensuite la suite de mesures d'occupations associées $(\nu_n)_{n \geq 1}$ par la formule suivante

$$\nu_n(dx) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \delta_{X_k}(dx)$$

ce qui conduit naturellement à la formule de récurrence suivante

$$\nu_{n+1} = \nu_n + \frac{\gamma_{n+1}}{\Gamma_n} (\delta_{X_{n+1}} - \nu_n).$$

Cette forme offre en particulier une définition via un algorithme stochastique de pas $(\frac{\gamma_{n+1}}{\Gamma_n})_{n \geq 1}$ ce qui rend l'implémentation numérique aisée. Le coup de calcul est de l'ordre de n^2 où n est l'horizon de temps.

Avant de présenter nos résultats, il convient d'être capable d'être capable de caractériser la mesure

invariante. Pour cela, introduisons la fonctionnelle d'énergie \mathcal{F} définie comme suit

$$\mathcal{F}(\nu) = \underbrace{\int_{\mathbb{R}^d} V(x)\nu(x)dx + \int_{\mathbb{R}^d} \nu(x) \log(\nu(x))dx}_{\text{Entropie}} + \underbrace{\frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x,y)\nu(x)\nu(y)dxdy}_{\text{Energie potentiel}}$$

Cette fonctionnelle traduit l'énergie du système associé à notre équation. Les lois de la Physique indiquent que $t \mapsto \mathcal{F}(\nu_t)$ est décroissante et converge vers son minimum. Un rapide calcul montre alors que les points critiques de \mathcal{F} sont les points fixes de l'application $\Pi : \nu \mapsto \Pi_\nu$ où Π_ν est la mesure invariante de (1.4.46).

Notre résultat va donc montrer que la suite $(\nu_n)_{n \geq 1}$ converge p.s vers un point fixe de l'application Π . Nous supposons d'ailleurs que ce point fixe est unique pour des raisons pratiques (unicité qui s'obtient en ajoutant des hypothèses sur V et W). Commençons par détailler les principales hypothèses entrant en jeu.

Hypothèses

$$(\mathbf{H_F}) \left\{ \begin{array}{l} \text{(i) } V \in C^2(\mathbb{R}^d) \text{ et } V \geq 0, W \in C^2(\mathbb{R}^{2d}) \text{ et } W \geq 0. \\ \text{(ii) Il existe une constante } M \text{ positive et telle que pour tout } x, y \in \mathbb{R}^d, \\ \qquad \qquad \qquad \langle \nabla V(x), \nabla_x W(x, y) \rangle \geq -M \\ \text{(iii) Il existe des constantes } a, b, c, d \text{ telles que pour tout } x, y \in \mathbb{R}^d, \\ \qquad \qquad \qquad \begin{array}{lcl} aV(x) - b & \leq & \|\nabla V(x)\|^2 \\ cW(x, y) - d & \leq & \|\nabla_x W(x, y)\|^2 \end{array} \\ \text{(iv) } \sup_{x \in \mathbb{R}^d} \|D^2 V(x)\| < \infty \text{ et } \sup_{x, y \in \mathbb{R}^d} \|D^2 W(x, y)\| < \infty. \\ \text{(v) Il existe des constantes positives } C_1, C_2 \text{ telles que pour tout } x, y \in \mathbb{R}^d \\ \qquad \qquad \qquad \|\nabla V(x)\|^2 \leq C_1 V(x) \quad \|\nabla_x W(x, y)\|^2 \leq C_2 W(x, y) \end{array} \right.$$

Ces hypothèses ont pour but d'assurer la tension de la suite de mesures d'occupation $(\nu_n)_{n \geq 1}$. Les hypothèses (i),(ii),(iii) et (iv) assurent à la fois la non-explosion du système et garantissent un dissipativité du système. L'hypothèse (v) permettra de caractériser la limite.

Introduisons les hypothèses relatives à l'opérateur Q_ν .

$$(\mathbf{H_Q}) \left\{ \begin{array}{l} \text{Pour toute fonction } f \in C_c^\infty(\mathbb{R}^d), \\ \text{(i) } x \mapsto Q_\nu f(x) \text{ est de classe } C^2(\mathbb{R}^d). \\ \text{(ii) } \sup_{x \in \mathbb{R}^d} |Q_\nu(f(x))| < \infty \\ \text{(iii) } \sup_{x \in \mathbb{R}^d} |\nabla Q_\nu(f(x))| < \infty \\ \text{(iv) } \sup_{x \in \mathbb{R}^d} |D_x^2 Q_\nu f(x)| < \infty \\ \text{(v) Pour toute mesure } \nu, \mu \in \mathcal{P}(\mathbb{R}^d), \\ |Q_\nu f(x) - Q_\mu f(x)| \leq C|f(x)|\|\nu - \mu\|_{TV} \end{array} \right.$$

(H_Q) est une hypothèse technique qui concerne la solution de l'équation de Poisson $f - \Pi_\nu(f) = L_\nu g$. Les conditions de régularités de Q_ν en la variable x sont classiques (lorsque la mesure ν est fixée). En revanche, la condition de régularité en la variable ν semble plus complexe à assurer dans la pratique et nous supposons cette hypothèse satisfaites en laissant cette question pour de futurs travaux. Pour une référence complète sur l'obtention des bornes apparaissant dans les hypothèses (i), (ii), (iii) et (iv), nous renvoyons le lecteur à l'ouvrage de Bakry, Gentil et Ledoux [13].

1.4.5 Résultat

Présentons notre résultat central.

THEOREME 1.4.20. *Supposons $(\mathbf{H_F})$, $(\mathbf{H_Q})$ satisfaites. Supposons de plus que V est fortement convexe, que W est symétrique et fortement convexe. Alors la suite $(\nu_n)_n$ converge presque sûrement vers ν^* , l'unique point fixe de l'application Π .*

Remarques L'unicité de la mesure invariante provient des hypothèses de forte convexité de V et W . On pourra s'interroger sur la nature du résultat si cette hypothèse est supprimée. De même, le choix du même pas pour l'algorithme et celui de la mesure d'occupation facilite notre travail mais on pourrait imaginer travailler avec des pas différents (voir [82]).

Chapitre 2

Regret bounds for Narendra and Shapiro bandit algorithms

Sommaire

2.1	Introduction	72
2.2	Definitions of the NS algorithms	75
2.2.1	Crude NSa and regret	75
2.2.2	Penalized and over-penalized two-armed NSa	76
2.2.3	Over-penalized multi-armed NSa	78
2.3	Main Results	78
2.3.1	Regret of the over-penalized two-armed bandit	78
2.3.2	Convergence of the multi-armed over-penalized bandit	80
2.3.3	Ergodicity of the limiting process	81
2.3.4	Total variation results	83
2.4	Proof of the regret bound (Theorems 2.3.1 and 2.3.2)	84
2.4.1	Notations	84
2.4.2	Evolution of $(Y_n)_{n \geq 1}$	85
2.4.3	Increase of exponent	87
2.4.4	Bound for $(\mathbb{E}(Z_n^{(3)}))_{n \geq n_0}$	90
2.4.5	Proof of Theorems 2.3.1 and 2.3.2 (a)	96
2.5	Almost sure and weak limit of the over-penalized bandit	98
2.5.1	A.s. convergence of the multi-armed bandit (Proposition 2.3.1)	98
2.5.2	Weak convergence of the normalized bandit (Proposition 2.3.2)	100
2.6	Ergodicity of the PDMP	102
2.6.1	Wasserstein results	102
2.6.2	Proof of total variation results	104
2.7	Technical result for the pseudo-regret upper bound	108

Abstract Narendra-Shapiro (NS) algorithms are bandit-type algorithms developed in the 1960s which have been deeply studied in infinite horizon but for which scarce non-asymptotic results exist. In this paper, we focus on a non-asymptotic study of the *regret* and address the following question : are Narendra-Shapiro bandit algorithms competitive from this point of view ? In our main result, we obtain some uniform explicit bounds for the regret of *(over)-penalized-NS* algorithms.

We also extend to the multi-armed case some convergence properties of penalized-NS algorithms towards a stationary Piecewise Deterministic Markov Process (PDMP). Finally, we establish some new sharp mixing bounds for these processes.

2.1 Introduction

The so-called Narendra-Shapiro bandit algorithm (referred to as NSa) was introduced in [98] and developed in [94] as a linear learning automata. This algorithm has been primarily considered by the probabilistic community as an interesting benchmark of stochastic algorithm. More precisely, NSa is an example of recursive (non-homogeneous) Markovian algorithm, topic whose almost complete historical overview may be found in the seminal contributions of [44] and [71].

NSa belongs to the large class of bandit-type policies whose principle may be sketched as follows : a d -armed bandit algorithm is a procedure designed to determine which one, among d sources, is the most profitable without spending too much time on the wrong ones. In the simplest case, the sources (or arms) randomly provide some rewards whose values belong to $\{0; 1\}$ with Bernoulli laws. The associated probabilities of success (p_1, \dots, p_d) are unknown to the player and his goal is to determine the most efficient source, *i.e.* the highest probability of success.

Let us now remind a rigorous definition of admissible sequential policies. We consider d independent sequences $(A_n^i)_{n \geq 0}$ of *i.i.d.* Bernoulli random variables $\mathcal{B}(p_i)$. Each A_n^i represents the reward associated with the arm i at time n . We then consider some sequential predictions where at each stage n a forecaster chooses an arm I_n , receives a reward $A_n^{I_n}$ and then uses this information to choose the next arm at step $n + 1$. As introduced in the pioneering work [111], the rewards are sampled independently of a fixed product distribution at each step n . The innovations here at time n are provided by $(I_n, A_n^{I_n})$ and we are naturally led to introduce the filtration $(\mathcal{F}_n)_{n \geq 0} := \left(\sigma((I_1, A_1^{I_1}), \dots, (I_n, A_n^{I_n})) \right)_{n \geq 0}$. In the following, the sequential admissible policies will be a $(\mathcal{F}_n)_{n \geq 0}$ (inhomogeneous) Markov chain. We also define another filtration by adding all the events before step n and observe that $(\bar{\mathcal{F}}_n)_{n \geq 0} := (\sigma((I_1, (A_1^j)_{1 \leq j \leq d}), \dots, (I_n, (A_n^j)_{1 \leq j \leq d})))_{n \geq 0}$. To sum-up, $\bar{\mathcal{F}}_n$ contains all the results of each arm between time 1 and n although \mathcal{F}_n only provides partial information about the tested arms.

In this paper, we focus on the stochastic NSa whose principle is very simple : it consists in sampling one arm according to a probability distribution on $\{1, \dots, d\}$, and in modifying this probability distribution in terms of the reward obtained with the chosen arm. From this point of view, this algorithm bears similarities with the EXP3 algorithm (and many of its variants) introduced in [6]. Among other close bandit algorithms, one can also cite the Thompson Sampling strategy where the random selection of the arm is based on a Bayesian posterior which is updated after each result. We refer to [1] for a recent theoretical contribution on this algorithm.

Instead of sampling one arm sequentially according to a randomized decision, other algorithms define their policy through a deterministic maximization procedure at each iteration. Among them, we can mention the UCB algorithm [5] and its derivatives (including MOSS [3] and KL-UCB [37]),

whose dynamics are dictated by an appropriate empirical upper confidence bound of the estimated best performance.

Let us now present the NSa algorithm. In fact, we will distinguish two types of NSa : crude-NSa and penalized-NSa. Before going further, let us recall their mechanism in the case of $d = 2$ (the general case will be introduced in Section 2.2). Designating X_n as the probability of drawing arm 1 at step n and $(\gamma_n)_{n \geq 0}$ as a decreasing sequence of positive numbers that tends to 0 when n goes to infinity, crude-NS is recursively defined by :

$$X_{n+1} = X_n + \begin{cases} \gamma_{n+1}(1 - X_n) & \text{if arm 1 is selected and wins} \\ -\gamma_{n+1}X_n & \text{if another arm is selected and wins} \\ 0 & \text{otherwise} \end{cases} \quad (2.1.1)$$

Note that the construction is certainly symmetric, *i.e.*, $1 - X_n$ (which corresponds to the probability of drawing arm 2) has a symmetric dynamics. The long-time behavior of some NSa was extensively investigated in the last decade. To name a few, in [78] and [76], some convergence and rate of convergence results are proved. However, these results strongly depend on both (γ_n) and the probabilities of success of the arms. In order to get rid of these constraints, the authors then introduced in [79] a penalized NSa and proved that this method is an efficient distribution-free procedure, meaning that it unconditionally converges to the best arm on the unknown probabilities p_1 and p_2 . The idea of the penalized-NS algorithm is to also take the failures of the player into account and to reduce the probability of drawing the tested arm when it loses. Designating $(\rho_n)_{n \geq 0}$ as a second positive sequence, the dynamics of the penalized NSa is given by :

$$X_{n+1} = X_n + \begin{cases} \gamma_{n+1}(1 - X_n) & \text{if arm 1 is selected and wins} \\ -\gamma_{n+1}X_n & \text{if arm 2 is selected and wins} \\ -\rho_n\gamma_{n+1}X_n & \text{if arm 1 is selected and loses} \\ \rho_{n+1}\gamma_{n+1}(1 - X_n) & \text{if arm 2 is selected and loses.} \end{cases} \quad (2.1.2)$$

Performances of bandit algorithms. In view of potential applications, it is certainly important to have some informations about the performances of the used policies. To this end, one first needs to define what is a “good” sequential algorithm. The primary efficiency requirement is the ability of the algorithm to asymptotically recover the best arm. In [79], this property is referred to as the *infallibility* of the algorithm. If without loss of generality, the first arm is assumed to be the best, (*i.e.* that $p_1 > \max\{p_2, \dots, p_d\}$) and if $X_n^{(1)}$ denotes the probability of drawing arm 1, the algorithm is said to be infallible if

$$\mathbb{P}(X_n^{(1)} \xrightarrow{n \rightarrow +\infty} 1) = 1. \quad (2.1.3)$$

An alternative way for describing the efficiency of a method is to consider the behaviour of the cumulative reward S_n obtained between time 1 and n :

$$S_n := \sum_{k=1}^n A_k^{I_k}.$$

In particular, in the old paper [11], Robbins is looking for algorithms such that

$$p_1 - \frac{\mathbb{E}[S_n]}{n} \xrightarrow{n \rightarrow +\infty} 0.$$

This last property is weaker than the infallibility of an algorithm since the Lebesgue theorem associated to (2.1.3) implies the convergence above.

A much stronger requirement involves the *regret* of the algorithm. The regret measures the gap between the cumulative reward of the best player and the one induced by the policy. The regret R_n is the \mathcal{F}_n -measurable random variable defined as :

$$R_n := \max_{1 \leq j \leq d} \sum_{k=1}^n [A_k^j - A_k^{I_k}]. \quad (2.1.4)$$

A good strategy corresponds to a selection procedure that minimizes the expected regret $\mathbb{E}R_n$, optimal ones being referred to as *minimax* strategies.

The former expected regret cannot be easily handled and is generally replaced in statistical analysis by the *pseudo-regret* defined as

$$\bar{R}_n := \max_{1 \leq j \leq d} \mathbb{E} \sum_{k=1}^n [A_k^j - A_k^{I_k}]. \quad (2.1.5)$$

Since $p_1 > p_j, \forall j \neq 1$, \bar{R}_n can also be written as

$$\bar{R}_n = \sum_{k=1}^n \mathbb{E}(A_k^1) - \mathbb{E} \left(\sum_{k=1}^n A_k^{I_k} \right) = n \left(p_1 - \frac{\mathbb{E}[S_n]}{n} \right).$$

A low *pseudo-regret* property then means that the quantity

$$n \left(p_1 - \frac{\mathbb{E}[S_n]}{n} \right)$$

has to be small, in particular sub-linear with n . The quantities R_n and \bar{R}_n are closely related and it is reasonable to study the pseudo-regret instead of the true regret, owing to the next proposition :

Proposition 2.1.1. (i) For any $(\mathcal{F}_n)_{n \geq 0}$ -measurable strategy, we obtain after n plays :

$$0 \leq \mathbb{E}R_n - \bar{R}_n \leq \sqrt{\frac{n \log d}{2}}.$$

(ii) Furthermore, for every integer n and d and for any (admissible) strategy,

$$\sup_{p_1 > p_2 \geq \dots \geq p_d} \mathbb{E}[R_n] \geq \frac{1}{20} \sqrt{nd}.$$

We refer to Proposition 34 of [4] for a detailed proof of (i) and to Theorem 5.1 of [6] for (ii). As mentioned in (ii), the bounds are distribution-free (uniform in p).¹ Since the MOSS method of [3]

1. The rate orders are strongly different if a dependence in p is allowed.

satisfies $\bar{R}_n \leq 25\sqrt{nd}$, (i) and (ii) show that a non-asymptotic distribution-free minimax rate is on the order of \sqrt{n} .

In particular, a fallible algorithm (meaning that $\mathbb{P}(X_n \xrightarrow{n \rightarrow +\infty} 1) < 1$) necessarily generates a linear regret and is not optimal. For example, in the case $d = 2$, the dependence of \bar{R}_n in terms of (X_n) is as follows :

$$\begin{aligned} \bar{R}_n &= p_1 n - \sum_{k=1}^n (p_1 \mathbb{E}[X_k] + p_2 \mathbb{E}[1 - X_k]) = (p_1 - p_2) \mathbb{E} \left[\sum_{k=1}^n (1 - X_k) \right] \\ &\gtrsim (p_1 - p_2) \mathbb{P}(X_\infty = 0) \times n. \end{aligned} \quad (2.1.6)$$

Objectives. In this paper, we therefore propose to focus on the regret and to answer to the question “Are NSa competitive from a regret viewpoint? In the case of positive answer, what are the associated upper-bounds?”

Due to some too restrictive conditions of infallibility, it will be seen that the crude-NSa cannot be competitive from a regret point of view. As mentioned before, the penalized NSa is more robust and is *a priori* more appropriate for this problem. More precisely, the penalty induces more balance between exploration and exploitation, *i.e.* between playing the best arm (the one in terms of the past actions) and exploring new options (playing the suboptimal arms). In this paper, we are going to prove that, up to a slight reinforcement, it is possible to obtain some competitive bounds for the regret of this procedure. The slightly modified penalized algorithm will be referred to as the *over-penalized*-algorithm below.

Outline. The paper is organized as follows : Section 2.2.1 provides some basic information about the crude NSa. Then, in Section 2.2.2, after some background on the penalized Nsa, we introduce a new algorithm called over-penalized NSa.

Section 2.3 is devoted to the main results : in Theorem 2.3.2, we establish an upper-bound of the *pseudo-regret* \bar{R}_n for the over-penalized algorithm in the two-armed case and also show a weaker result for the penalized NSa.

In this section, we also extend to the multi-armed case some existing convergence and rate of convergence results of the two-armed algorithm. In the “critical” case (see below for details), the normalized algorithm converges in distribution toward a PDMP (Piecewise Deterministic Markov Process). We develop a careful study of its ergodicity and bounds on the rate of convergence to equilibrium are established. It uses a non-trivial coupling strategy to derive explicit rates of convergence in Wasserstein and total variation distance. The dependence of these rates are made explicit with the several parameters of the initial Bandit problem.

The rest of the paper is devoted to the proofs of the main results : Section 2.4 is dedicated to the regret analysis, and Section 2.5 establishes the weak limit of the rescaled multi-armed bandit algorithm. Finally, Section 2.6 includes all the proofs of the ergodic rates.

2.2 Definitions of the NS algorithms

2.2.1 Crude NSa and regret

The crude NSa (2.1.1) is rather simple : it defines a $(\mathcal{F}_n)_{n \geq 0}$ Markov chain $(X_n)_{n \geq 0}$ and I_n is a random variable satisfying :

$$\mathbb{P}(I_{n+1} = 1 | \mathcal{F}_n) = X_n \quad \text{and} \quad \mathbb{P}(I_{n+1} = 2 | \mathcal{F}_n) = 1 - X_n$$

The arm I_{n+1} is selected at step $n+1$ with the current distribution $(X_n, 1 - X_n)$ and is evaluated. In the event of success, the weight of the arm I_{n+1} is increased and the weight of the other arm is decreased by the same quantity. The algorithm can be rewritten in a more concise form as :

$$X_{n+1} = X_n + \gamma_{n+1}(1_{I_{n+1}=1} - X_n)A_{n+1}^{I_{n+1}}. \quad (2.2.7)$$

The arm i at step n succeeds with the probability $p_i = \mathbb{P}(A_n^i = 1)$ and we suppose *w.l.o.g.* that $p_1 > p_2$ so that the arm 1 is the optimal one.

As pointed in (2.1.6), we obtain that

$$\bar{R}_n = (p_1 - p_2)\mathbb{E} \left[\sum_{k=1}^n (1 - X_k) \right].$$

This formula is important regarding the fallibility of an algorithm. In particular, it is shown in [76] that for any choice $\gamma_n = C(n+1)^{-\alpha}$ with $\alpha \in (0, 1)$ and $C > 0$ or $\gamma_n = C/(n+1)$ with $C > 1$, the NSa (2.2.7) may be fallible : some parameters (p_1, p_2) exist such that $(X_n)_{n \geq 0}$ a.s. converges to a binary random variable X_∞ with $\mathbb{P}(X_\infty = 0) > 0$. In this situation, for large enough n , we have :

$$\bar{R}_n \gtrsim (p_1 - p_2)\mathbb{P}(X_\infty = 0) \times n \gg \sqrt{n}$$

It can easily be concluded that this method cannot induce a competitive policy since some “bad” values of the probabilities (p_1, p_2) generate a linear regret.

2.2.2 Penalized and over-penalized two-armed NSa

Penalized NSa. A major difference between the crude NSa and its penalized counterpart introduced in [79] relies on the exploitation of the failure of the selected arms. The crude NSa (2.1.1) only uses the sequence of successes to update the probability distribution $(X_n, 1 - X_n)$ since the value of X_n is modified *iff* $A_n^{I_n} = 1$. In contrast, the penalized NSa (2.1.2) also uses the information generated by a potential failure of the arm I_{n+1} . More precisely, in the event of success of the selected arm I_{n+1} , this penalized NSa mimics the crude NSa, whereas in the case of failure, the weight of the selected arm is now multiplied (and thus decreased) by a factor $(1 - \gamma_{n+1}\rho_{n+1})$ (whereas the probability of drawing the other arm is increased by the corresponding quantity). For the penalized NSa, the update formula of $(X_n)_{n \geq 1}$ can be written in the following way :

$$\begin{aligned} X_{n+1} &= X_n + \gamma_{n+1} [1_{I_{n+1}=1} - X_n] A_{n+1}^{I_{n+1}} \\ &\quad - \gamma_{n+1}\rho_{n+1} [X_n 1_{I_{n+1}=1} - (1 - X_n) 1_{I_{n+1}=2}] (1 - A_{n+1}^{I_{n+1}}). \end{aligned} \quad (2.2.8)$$

Over-penalized NSa. In view of the minimization of the regret, we will show that it may be useful to reinforce the penalization. For this purpose, we introduce a slightly “over-penalized” NSa where a player is also (slightly) penalized if it wins :

- If the arm 1 wins, then with probability $1 - \sigma$ it is penalized by a factor $\gamma_{n+1}\rho_{n+1}X_n$.
- If the arm 2 wins, then with probability $1 - \sigma$ arm 1 is increased by a factor of $\gamma_{n+1}\rho_{n+1}(1 - X_n)$.

The over-penalized-NSa can be written as follows

$$X_{n+1}^\sigma = X_n^\sigma + \gamma_{n+1} [1_{I_{n+1}=1} - X_n^\sigma] A_{n+1}^{I_{n+1}} - \gamma_{n+1} \rho_{n+1} [X_n^\sigma 1_{I_{n+1}=1} - (1 - X_n^\sigma) 1_{I_{n+1}=2}] \left(1 - A_{n+1}^{I_{n+1}} B_{n+1}^\sigma\right) \quad (2.2.9)$$

where $(B_n^\sigma)_n$ is a sequence of i.i.d. r.v. with a Bernoulli distribution $\mathcal{B}(\sigma)$, meaning that $\mathbb{P}(B_n^\sigma = 0) = 1 - \sigma$. Moreover, these r.v. are independent of $(A_n^j)_{n,j}$ and in such a way that for all $n \in \mathbb{N}$, B_n^σ and I_n are also independent. It should be noted that

$$1 - A_n^{I_n} B_n^\sigma = [1 - A_n^{I_n}] + A_n^{I_n} (1 - B_n^\sigma).$$

In fact, this slight over-penalization of the successful arm (with probability σ) can be viewed as an additional statistical excitation which helps the stochastic algorithm to escape from local traps. The case $\sigma = 1$ corresponds to the penalized NSa (2.2.8), whereas when $\sigma = 0$, the arm is always penalized when it plays. In particular, this modification implies that the increment of X_n^σ is slightly weaker than in the previous case when the selected arm wins.

Asymptotic convergence of the penalized NSa. Before stating the main results, we need to understand which regret \bar{R}_n could be reached by penalized and over-penalized NSa. We recall (in a slightly less general form) the convergence results of Proposition 3, Theorems 3 and 4 of [79].

Theorem 2.2.1 (Lamberton & Pages, [79]). Let $0 \leq p_2 < p_1 \leq 1$ and $\gamma_n = \gamma_1 n^{-\alpha}$ and $\rho_n = \rho_1 n^{-\beta}$ with $(\alpha, \beta) \in (0, +\infty)$ and $(\gamma_1, \rho_1) \in (0, 1)^2$. Let $(X_n)_n$ be the algorithm given by (2.2.8).

i) If $0 < \beta \leq \alpha$ and $\alpha + \beta \leq 1$, the penalized two-armed bandit is *infallible*.

ii) Furthermore, if $0 < \beta < \alpha$ and $\alpha + \beta < 1$, then $\frac{1 - X_n}{\rho_n} \xrightarrow{\text{a.s.}} \frac{1 - p_1}{p_1 - p_2}$.

iii) If $\alpha = \beta \leq 1/2$ and $g = \gamma_1/\rho_1 : \frac{1 - X_n}{\rho_n} \xrightarrow{w^*} \mu$, where $\xrightarrow{w^*}$ stands for the convergence in distribution and μ is the stationary distribution of the PDMP whose generator \mathcal{L} acts on $\mathcal{C}_c^1(\mathbb{R}_+)$ as

$$\forall f \in \mathcal{C}_c^1(\mathbb{R}_+) \quad \mathcal{L}f(y) = p_2 y \frac{f(y+g) - f(y)}{g} + (1 - p_1 - p_1 y) f'(y).$$

In view of Theorem 2.2.1, we can use formula (2.1.6) to obtain

$$\bar{R}_n = (p_1 - p_2) \sum_{k=1}^n \rho_k \mathbb{E} \left(\frac{1 - X_k}{\rho_k} \right). \quad (2.2.10)$$

We then obtain the key observation

$$\sup_{n \in \mathbb{N}} \mathbb{E} \left[\frac{1 - X_n}{\rho_n} \right] \leq C < +\infty \implies \bar{R}_n \leq C(p_1 - p_2) \sum_{k=1}^n \rho_k, \quad (2.2.11)$$

where C is a constant that may depend on p_1 and p_2 . According to Theorem 2.2.1, it seems that the potential optimal choice corresponds to the one of (iii). Indeed, the infallibility occurs only when $\alpha \geq \beta$ and $\alpha + \beta \leq 1$ and Equation (2.2.10) suggests that β should be chosen as large as

possible to minimize the r.h.s. of (2.2.11), leading to $\alpha = \beta = 1/2$. This is why in the following, we will focus on the case :

$$\gamma_n = \frac{\gamma_1}{\sqrt{n}} \quad \text{and} \quad \rho_n = \frac{\rho_1}{\sqrt{n}}. \quad (2.2.12)$$

2.2.3 Over-penalized multi-armed NSa

We generalize the definition of the penalized and over-penalized NSa to the d -armed case, with $d \geq 2$. Let $p = (p_1, \dots, p_d) \in (0, 1)^d$ and assume that $A_n^j \sim \mathcal{B}(p_j)$ (p_i the probability of success of arm i). The over-penalized NSa recursively defines a sequence of probability measures on $\{1, \dots, d\}$ denoted by $(\Pi_n)_{n \geq 1}$ where $\Pi_n = (X_n^1, \dots, X_n^d)$. At step n , the arm I_{n+1} is sampled according to the discrete distribution X_n and then tested through the computation of $A_{n+1}^{I_{n+1}}$. Setting $j \in \{1, \dots, d\}$, the multi-armed NSa is defined by :

$$\begin{aligned} X_{n+1}^j &= X_n^j + \gamma_{n+1} [1_{I_{n+1}=j} - X_n^j] A_{n+1}^{I_{n+1}} \\ &\quad - \gamma_{n+1} \rho_{n+1} X_n^{I_{n+1}} (1 - A_{n+1}^{I_{n+1}} B_{n+1}^\sigma) \left[1_{I_{n+1}=j} - \frac{1 - 1_{I_{n+1}=j}}{d-1} \right]. \end{aligned} \quad (2.2.13)$$

In contrast with the two-armed case, we have to choose how to distribute the penalty to the other arms when $d > 2$. The (natural) choice in (2.2.13) is to divide it fairly, *i.e.*, to spread it uniformly over the other arms. Note that alternative algorithms (not studied here) could be considered.

2.3 Main Results

2.3.1 Regret of the over-penalized two-armed bandit

First, we provide some uniform upper-bounds for the two-armed σ -over-penalized NSa. Our main result is Theorem 2.3.2. Before stating it, we choose to state a new result when $\sigma = 1$, *i.e.* for the “original” penalized NSa introduced in [79].

Theorem 2.3.1. Let $(X_n)_{n \geq 0}$ be the two-armed penalized NSa defined by (2.2.8) with $(\gamma_n, \rho_n)_{n \geq 1}$ defined by (2.2.12) with $(\gamma_1, \rho_1) \in (0, 1)^2$. Then, for every $\delta \in (0, 1)$, a positive C_δ exists such that :

$$\forall n \in \mathbb{N}, \quad \sup_{(p_1, p_2) \in [0, 1], p_2 \leq p_1 \wedge (1-\delta)} \bar{R}_n \leq C_\delta \sqrt{n}.$$

Remark 2.3.1. The upper bound of the original penalized-NS algorithm is not completely uniform. From a theoretical point of view, there is not enough penalty when p_2 is too large, which in turn generates a deficiency of the mean-reverting effect for the sequence $((1 - X_n)/\rho_n)_{n \geq 1}$ when X_n is close to 0. In other words, the trap of the stochastic algorithm near 0 is not enough repulsive and Figure 2.1 below shows that this problem also appears numerically and suggests a logarithmic explosion of $\sup_{p_1 < p_2} \bar{R}_n / \sqrt{n}$.

This explains the interest of the over-penalization, illustrated by the next result, which is the main theorem of the paper.

Theorem 2.3.2. Let $(X_n)_{n \geq 0}$ be the two-armed σ -over-penalized NSa defined by (2.2.9) with $\sigma \in [0, 1)$ and $(\gamma_n, \rho_n)_{n \geq 1}$ defined by (2.2.12) with $(\gamma_1, \rho_1) \in (0, 1)^2$. Then,

(a) A $C_\sigma(\gamma_1, \rho_1)$ exists such that :

$$\forall n \in \mathbb{N}, \quad \sup_{(p_1, p_2) \in [0, 1], p_2 < p_1} \bar{R}_n \leq C_\sigma(\gamma_1, \rho_1) \sqrt{n}.$$

(b) Furthermore, the choice $\sigma = 0$, $\gamma_n = 2.63\rho_n = 0.89/\sqrt{n}$ yields

$$\forall n \in \mathbb{N}, \quad \sup_{(p_1, p_2) \in [0, 1], p_2 < p_1} \bar{R}_n \leq 31.1\sqrt{2n}. \quad (2.3.14)$$

Remark 2.3.2. At the price of technicalities, C_σ could be made explicit in terms of γ_1 and ρ_1 for every $\sigma > 0$. The second bound is obtained by an optimization of $C_0(\gamma_1, \rho_1)$ (see (2.4.38) and below).

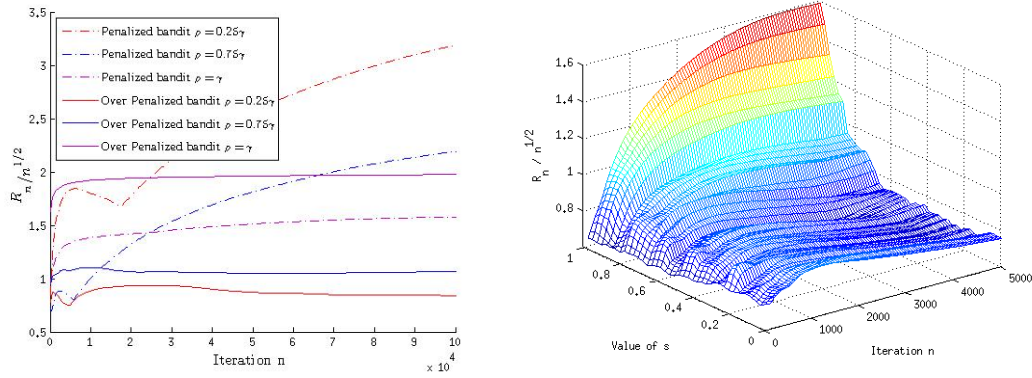


FIGURE 2.1 – Evolution of $n \mapsto \sup_{(p_1, p_2) \in [0, 1], p_2 \leq p_1} \frac{\bar{R}_n}{\sqrt{n}}$ for the over-penalized algorithm. Left : $\sigma = 0$, Right : σ varies from 0 to 1.

Figure 2.1 presents on the left side a numerical approximation of $n \mapsto \sup_{p_2 < p_1} \bar{R}_n / \sqrt{n}$ for the penalized and over-penalized algorithms. The continuous curves indicate that the upper bound $31.1\sqrt{2}$ in Theorem 2.3.2 is not sharp since the over-penalized NSa satisfies a uniform upper-bound on the order of $0.9\sqrt{n}$. This bound is obtained with a small σ (as pointed in Theorem 2.3.2), and $\gamma_n = \frac{1}{\sqrt{4+n}} = 4\rho_n$ (red line in Figure 2.1 (left)), suggesting that the rewards should *always* be over-penalized with $\rho_n = \frac{\gamma_n}{4}$.

The right-hand side of Figure 2.1 focuses on the behavior of the regret with σ . The map $(n, \sigma) \mapsto \sup_{p_1 < p_2} \bar{R}_n / \sqrt{n}$ confirms the influence of the over-penalization and indicates that to obtain optimal performances for the cumulative regret, we should use a low value of σ between 0 and $3/5$. The importance of this choice of σ seems relative since the behaviour of the over-penalized bandit is stable on this interval. The best numerical choice is attained for $\sigma = 1/4$ and $\rho_n = \frac{1}{4}\gamma_n$ and permits to achieve a long-time behavior of \bar{R}_n / \sqrt{n} of the order $3/4$ (see Figure 2.2, red line).

Finally, the statistical performances of the over-penalized NSa are compared with some classical bandit algorithms : KL-UCB algorithm (see *e.g.* [37] and the references therein) and EXP3 (see [6]). These two algorithms are anytime policies that are known to be minimax optimal with a cumulative minimax regret of the order \sqrt{n} . Figure 2.2 shows that the performances of the over-

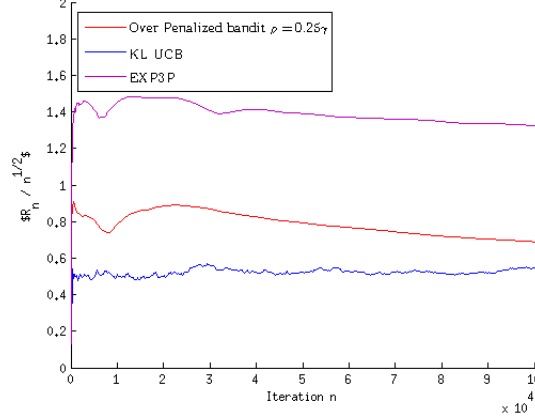


FIGURE 2.2 – Evolution of $n \mapsto \sup_{(p_1, p_2) \in [0, 1], p_2 \leq p_1} \frac{\bar{R}_n}{\sqrt{n}}$ for the over-penalized algorithm (with $\sigma = \frac{1}{4}$) and comparison with EXP3 and KL-UCB.

penalized NSa are located between the one of the KL-UCB² algorithm and of the EXP3 algorithm (our simulations suggest that the uniform bounds of KL-UCB and EXP3 are respectively 1/2 and 3/2). Also, it is worth noting that the simulation cost of the over-penalized NSa is strongly weaker than the initial UCB algorithm (the phenomenon is increased when compared to KL-UCB, which requires an additional difficulty for the computation of the upper confidence bound at each step) : the same amount of Monte-Carlo simulations for the over-penalized NSa is almost hundred times faster than the KL-UCB runs in equivalent numerical conditions.

2.3.2 Convergence of the multi-armed over-penalized bandit

We first extend Theorem 2.2.1 of [79] to the over-penalized NSa in the multi-armed situation. The result describes the pointwise convergence.

Proposition 2.3.1 (Convergence of the multi-armed over-penalized bandit). Consider $p_d \leq \dots \leq p_2 < p_1$ and $\gamma_n = \gamma_1 n^{-\alpha}$, $\rho_n = \rho_1 n^{-\beta}$ with $(\alpha, \beta) \in (0, +\infty)$ and $(\gamma_1, \rho_1) \in (0, 1)^2$. Algorithm (2.2.8) with $\sigma \in (0, 1]$ satisfies

- i) If $0 < \beta \leq \alpha$ and $\alpha + \beta \leq 1$, then $\lim_{n \rightarrow +\infty} \Pi_n = (1, 0, \dots, 0)$ a.s.
- ii) Furthermore, if $0 < \beta < \alpha$ and $\alpha + \beta < 1$, then :

$$\forall i \in \{2, \dots, d\}, \quad \frac{X_n^i}{\rho_n} \longrightarrow \frac{1 - \sigma p_1}{(d-1)(p_1 - p_i)} \quad \text{a.s.}$$

Proposition 2.3.2 provides a description of the behavior of the *normalized* NSa while considering $Y_n^j = \frac{X_n^j}{\rho_n}$. It states that $(Y_n)_{n \geq 0}$ converges to the dynamics of a *Piecewise Deterministic Markov Process* (referred to as PDMP below).

2. KL-UCB is performed with the Matlab package available on the website <http://mloss.org/software/view/415/>

Proposition 2.3.2 (Weak convergence of the over-penalized NSa). Under the assumptions of Proposition 2.3.1, if $\alpha = \beta \leq 1/2$ and $g = \gamma_1/\rho_1$, then :

$$\frac{1}{\rho_n} (X_n^2, \dots, X_n^d) \xrightarrow{w^*} \mu_d,$$

where μ_d is the (unique) stationary distribution of the Markov process whose generator \mathcal{L}_d acts on compactly supported functions f of $\mathcal{C}^1((\mathbb{R}_+)^{d-1})$ as follows :

$$\begin{aligned} \mathcal{L}_d f(y_2, \dots, y_d) &= \sum_{i=2, \dots, d} \frac{p_i y_i}{g} (f(y_2, \dots, y_i + g, \dots, y_d) - f(y_2, \dots, y_i, \dots, y_d)) \\ &+ \sum_{i=2, \dots, d} \left(\frac{1 - \sigma p_1}{d-1} - p_1 y_i \right) \partial_i f(y_2, \dots, y_d). \end{aligned} \quad (2.3.15)$$

2.3.3 Ergodicity of the limiting process

In this section, we focus on the long time behavior of the limiting Markov process that appears (after normalization) in Proposition 2.3.2. As mentioned before, this process is a PDMP and its long time behavior can be carefully studied with some arguments in the spirit of [15]. We also learned about the existence of a close study in the PhD thesis of Florian Bouguet (some details may be found in [31]). Such properties are stated for both the one-dimensional and the multidimensional cases.

One-dimensional case

Setting

$$a = 1 - p_1, \quad b = p_1, \quad g = \frac{\gamma_1}{\rho_1}, \quad c = \frac{p_2}{g},$$

the generator \mathcal{L} given by Proposition 2.3.2 may be written as :

$$\forall f \in \mathcal{C}^1(\mathbb{R}_+^*, \mathbb{R}), \quad \mathcal{L}f(x) = \underbrace{(a - bx)f'(x)}_{\text{deterministic part}} + \underbrace{cx}_{\text{jump rate}} \underbrace{(f(x+g) - f(x))}_{\text{jump size}}. \quad (2.3.16)$$

In what follows, we will assume that a, b, c and g are positive numbers. We can see in \mathcal{L} two parts. On the one hand, the deterministic flow that guides the PDMP between the jumps is given by :

$$\begin{cases} \partial_t \phi(x, t) &= (a - bx) \partial_x \phi(x, t) \\ \phi(x, 0) &= x \in \mathbb{R}_+^* \end{cases}$$

so that

$$\phi(x, t) = \frac{a}{b} + \left(x - \frac{a}{b} \right) e^{-bt}.$$

Hence, if $x > \frac{a}{b}$ (resp. $x < \frac{a}{b}$), $t \mapsto \phi(x, t)$ decreases (resp. increases) and converges exponentially fast to $\frac{a}{b}$.

On the other hand, the PDMP possesses some positive jumps that occur with a Poisson intensity “ cx ”, whose size is deterministic and equals to g .

From the finiteness and positivity of g , it is easy to show that for every positive starting point, the process is *a.s.* well-defined on \mathbb{R}_+ , positive and does not explode in finite time. The fact that the

size of the jumps is deterministic is less important and what follows could easily be generalized to a random size g (under adapted integrability assumptions). In Figure 2.3 below, some paths of the process are represented with different values of the parameters.

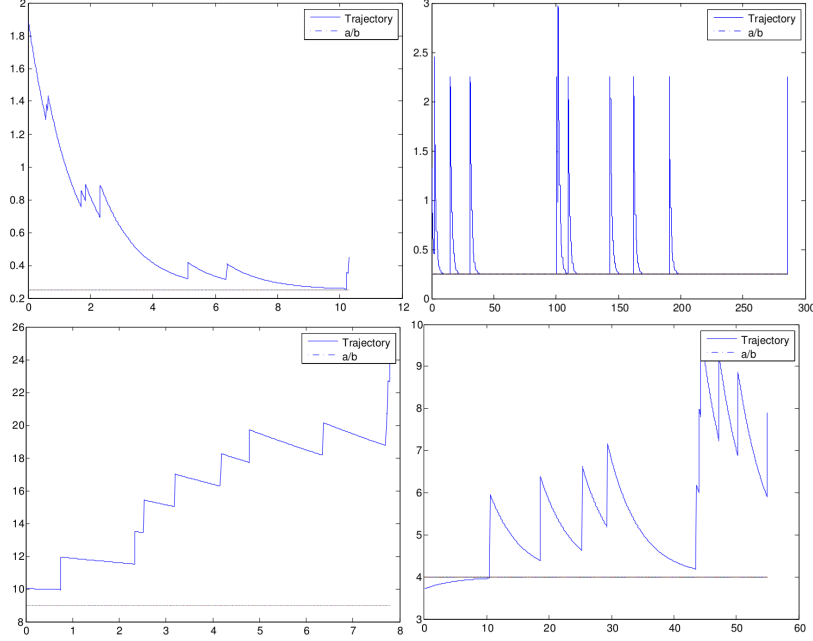


FIGURE 2.3 – Exact simulation of trajectories of a process driven by (2.3.16) when $g = 0.1, a = 0.2, b = 0.8, c = 0.2$ (top left) $g = 2, a = 0.2, b = 0.8, c = 0.1$ (top right), $g = 2, a = 0.9, b = 0.9, c = 0.15$ (bottom left) and $g = 2, a = 0.8, b = 0.2, c = 0.05$ (bottom right).

Convergence results

As pointed out in Figure 2.3, the long-time behavior of the process certainly depends on the relationship between the mean-reverting effect generated by “ $-bx$ ” and the frequency and size of the jumps.

Invariant measure The process (2.3.16) possesses a unique invariant distribution if $b - cg > 0$. Actually, the existence is ensured by the fact that $V(x) = x$ is a Lyapunov function for the process since

$$\forall x \in \mathbb{R}_+^*, \quad \mathcal{L}V(x) = a - (b - cg)x = a - (b - cg)V(x)$$

Among other arguments, the uniqueness is ensured by Theorem 2.3.3 (the convergence in Wasserstein distance of the process toward the invariant distribution implies in particular its uniqueness). We denote it by μ_∞ below. It could also be shown that $\text{Supp}(\mu_\infty) = (a/b, +\infty)$, that the process is strongly ergodic on $(a/b, +\infty)$ (see [48] for some background) and that if $b - cg > 0$, the process explodes when $t \rightarrow +\infty$ (this case corresponds to the bottom left-hand side of Figure 2.3). Finally,

it should be noted that for the limiting PDMP of the bandit algorithm,

$$b - cg = p_1 - p_2 = \pi$$

and thus, the ergodicity condition coincides with the positivity of π .

Wasserstein results We aim to derive rates of convergence for the PDMP toward μ_∞ for two distances, namely the Wasserstein distance and the total variation distance. Rather different ways to obtain such results exist using coupling arguments or PDEs. We use coupling techniques here that are consistent with the work of [15] and [41]. Before stating our results, let us recall that the p -Wasserstein distance is defined for any probability measures μ and ν on \mathbb{R}^d by :

$$\mathcal{W}_p(\mu, \nu) = \inf \left\{ \mathbb{E}(|X - Y|^p)^{\frac{1}{p}} \mid \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu \right\},$$

where $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ respectively denote the laws of X and Y . Designating μ_0 as the initial distribution of the PDMP and μ_t as its law at time t , we now state the main result on the PDMP in dimension one driven by (2.3.16).

Theorem 2.3.3 (One dimensional PDMP). Let $p \geq 1$ and denote for every $t \geq 0$ $\mu_t := \mathcal{L}(X_t^{\mu_0})$ where $(X_t^{\mu_0})$ is a Markov process driven by (2.3.16) with initial distribution μ_0 (with support included in \mathbb{R}_+^*). If $p = 1$, we have

$$\left| \int x(\mu_0 - \mu_\infty)(dx) \right| e^{-\pi t} \leq \mathcal{W}_1(\mu_t, \mu_\infty) \leq \mathcal{W}_1(\mu_0, \mu_\infty) e^{-t\pi}$$

and if $p > 1$, a constant γ_p exists such that

$$\mathcal{W}_p(\mu_t, \mu_\infty) \leq \gamma_p e^{-\frac{t\pi}{p}}.$$

where $(\gamma_p)_{p \geq 1}$ satisfies the recursion $\gamma_p^p = \gamma_{p-1}^{p-1}[pa + (1+g)^p]$.

Remark 2.3.3. If $p = 1$, the lower and upper bounds imply the optimality of the rate obtained in the exponential. For $p > 1$, the optimality of the exponent $e^{-\pi t/p}$ is still an open question.

We now give a corollary for the limiting process that appears in Proposition 2.3.2.

Corollary 2.3.1 (Multi-dimensional PDMP). Let $(Y_t)_{t \geq 0}$ be the PDMP driven by (2.3.15) with initial distribution $\mu_0 \in (\mathbb{R}_+^*)^d$. Then, the conclusions of Theorem 2.3.3 hold with $\pi = p_1 - p_2$.

The proof is almost obvious due to the “tensorized” form of the generator \mathcal{L}_d . Actually, for every starting point $y = (y_2, \dots, y_d)$, all the coordinates $(Y_t^i)_{t \geq 0}$ are independent one-dimensional PDMPs with generator \mathcal{L} defined by (2.3.16) with

$$a_i = \frac{1 - \sigma p_1}{d - 1}, \quad b_i = p_1 \quad \text{and} \quad c_i = p_i/g. \quad (2.3.17)$$

The result then easily follows from Theorem 2.3.3 with a global rate given by $\min\{b_i - c_i g, i = 2, \dots, d\} = p_1 - p_2$. The details are left to the reader.

2.3.4 Total variation results

When some bounds are available for the Wasserstein distance, a classical way to deduce an upper bound of the total variation is to build a two-step coupling. In the first step, a Wasserstein coupling

is used to bring the paths sufficiently close (with a probability controlled by the Wasserstein bound). In a second step, we use a total variation coupling to try to stick the paths with a high probability. In our case, the jump size is deterministic and sticking the paths implies a non trivial coupling of the jump times. Some of the ideas to obtain the results below are in the spirit of [15], who follows this strategy for the TCP process.

Theorem 2.3.4. Denote by $(P_t)_{t \geq 0}$ the semi-group related to \mathcal{L} defined in (2.3.16) and let μ_0 be a starting distribution with moments of any order. Then, for every $\varepsilon > 0$, a $C_\varepsilon > 0$ exists such that :

$$\|\mu_0 P_t - \mu_\infty P_t\|_{TV} \leq C_\varepsilon e^{-(\alpha\pi - \varepsilon)t} \quad \text{with } \alpha = \frac{1}{2 + \frac{b\pi}{ac}}.$$

Once again, this result can be extended to the multi-armed case.

Corollary 2.3.2. Let $(Y_t)_{t \geq 0}$ be the PDMP driven by (2.3.15) with initial distribution $\mu_0 \in (\mathbb{R}_+^*)^d$. Then, the conclusions of Theorem 2.3.4 hold with $\alpha\pi$ replaced by :

$$\sum_{i=2}^d \frac{1}{2 + \frac{b_i \pi_i}{a_i c_i}} \pi_i$$

where $\pi_i = p_1 - p_i$ and a_i, b_i and c_i are defined by (2.3.17).

The proof of this result is based on the remark that follows Corollary 2.3.1. Owing to the “tensorization” property, the probability for coupling all the coordinates before time t is essentially the product of the probabilities of the coupling of each coordinate. Once again, the details of this corollary are left to the reader.

2.4 Proof of the regret bound (Theorems 2.3.1 and 2.3.2)

This section is devoted to the study of the regret of the penalized two-armed bandit procedure described in Section 2.1.2. We will mainly focus on the proof of the explicit bound given in Theorem 2.3.2(b) and we will give the main ideas for the proofs of Theorems 2.3.1 and 2.3.1(a).

2.4.1 Notations

In order to lighten the notations, X_n^1 will be summarized by X_n , so that $X_n^2 = 1 - X_n$.

The proofs are then strongly based on a detailed study of the behavior of the (positive) sequence $(Y_n)_{n \geq 1}$ defined by

$$\forall n \geq 1 \quad Y_n = \frac{1 - X_n}{\gamma_n}. \quad (2.4.18)$$

As we said before, we will consider the following sequences $(\gamma_n)_{n \geq 1}$ and $(\rho_n)_{n \geq 1}$ below :

$$\forall n \geq 1, \quad \gamma_n = \frac{\gamma_1}{\sqrt{n}} \quad \text{and} \quad \rho_n = \frac{\rho_1}{\sqrt{n}} = \tilde{\rho}_1 \gamma_n \quad \text{and} \quad \tilde{\rho}_1 = \frac{\rho_1}{\gamma_1},$$

where γ_1 and ρ_1 are constants in $(0, 1)$ that will be specified later. In the meantime, we also define :

$$\pi = p_1 - p_2 \in (0, 1).$$

With this setting, the pseudo-regret is

$$\bar{R}_n = \pi \sum_{k=1}^n \gamma_k \mathbb{E}[Y_k].$$

It should be noted here that we have substituted the division by ρ_k in (2.2.11) by a normalization with γ_k . This will be easier to handle in the sequel. The main issue now is to obtain a convenient upper bound for $\mathbb{E}[Y_k]$. More precisely, note that :

$$\forall n_0 \in \mathbb{N} \quad \forall n \leq n_0 - 1, \quad \bar{R}_n \leq \pi n \leq \pi \sqrt{n_0 - 1} \sqrt{n},$$

and conversely for every $n \geq n_0$,

$$\begin{aligned} \frac{\bar{R}_n}{\sqrt{n}} &\leq \pi \sqrt{n_0 - 1} + \pi \sup_{n \geq n_0} \mathbb{E}[Y_n] \frac{1}{\sqrt{n}} \sum_{n=n_0}^n \frac{\gamma_1}{\sqrt{k}} \\ &\leq \pi \left(\sqrt{n_0 - 1} + 2\gamma_1 \sup_{n \geq n_0} \mathbb{E}[Y_n] \right). \end{aligned} \quad (2.4.19)$$

Thus it is enough to derive an upper bound of $\mathbb{E}[Y_n]$ after an iteration n_0 that can be on the order of $1/\pi^2$. In particular, the “suitable” choice of n_0 will strongly depend on the value of π .

2.4.2 Evolution of $(Y_n)_{n \geq 1}$

Recursive dynamics of $(Y_n)_{n \geq 1}$. In order to understand the mechanism and difficulties of the penalized procedure, let us first roughly describe the behavior of the sequences $(X_n)_{n \geq 1}$ and $(Y_n)_{n \geq 1}$. According to (2.2.9),

$$\begin{aligned} \mathbb{E}[X_{n+1} | \mathcal{F}_n] &= X_n + \gamma_{n+1} X_n (1 - X_n) [p_1 - p_2] \\ &\quad + \gamma_{n+1} \rho_{n+1} [(1 - X_n)^2 (1 - \sigma p_2) - X_n^2 (1 - \sigma p_1)]. \end{aligned}$$

It can be observed that the drift term may be split into two parts, where the main part is the usual drift of NSa described by h defined by :

$$\forall x \in [0, 1], \quad h(x) = [p_1 - p_2]x(1 - x). \quad (2.4.20)$$

The second term comes from the penalization procedure and depends on σ . We set

$$\kappa_\sigma(x) = (1 - \sigma p_2)(1 - x)^2 - (1 - \sigma p_1)x^2. \quad (2.4.21)$$

As a consequence, we can write the evolution of $(X_n)_{n \geq 0}$ as follows :

$$1 - X_{n+1} = 1 - X_n - \gamma_{n+1} [h(X_n) + \rho_{n+1} \kappa_\sigma(X_n) + \Delta M_{n+1}], \quad (2.4.22)$$

where ΔM_{n+1} is a martingale increment. On the basis of the equation above, we easily derive that

$$\forall n \geq 1, \quad Y_{n+1} = Y_n (1 + \gamma_n (\epsilon_n - \pi X_n)) - \rho_{n+1} \kappa_\sigma(X_n) + \Delta M_{n+1}$$

where

$$\epsilon_n = \frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} = \frac{1}{\gamma_1} (\sqrt{n+1} - \sqrt{n}) \leq \frac{1}{2\gamma_1\sqrt{n}} = \frac{\gamma_n}{2\gamma_1^2}. \quad (2.4.23)$$

It follows that the increments of $(Y_n)_{n \geq 1}$ are given by :

$$\Delta Y_{n+1} := Y_{n+1} - Y_n = \gamma_n \varphi_n(Y_n) - \Delta M_{n+1}$$

where the drift function φ_n acting on the sequence $(Y_n)_{n \geq 1}$ is defined as

$$\varphi_n(y) = \underbrace{y \times [\epsilon_n + \pi(\gamma_n y - 1)]}_{:= \varphi_n^1(y)} + \underbrace{\left(-\frac{\rho_{n+1}}{\gamma_n} \kappa_\sigma(1 - \gamma_n y) \right)}_{:= \varphi_n^2(y)}.$$

To better understand the underlying effects of the dynamical system, it should be recalled that the definition of the sequence $(Y_n)_{n \geq 1}$ implies that $Y_n \in [0, \gamma_n^{-1}]$ with $\gamma_n^{-1} \sim Cn^{1/2}$. Since we aim to obtain a uniform bound (over n) of $\mathbb{E}[Y_n]$, it is thus important to understand the behavior of the drift φ_n over $[0, \gamma_n^{-1}]$. In particular, it is of primary interest to see where the function φ_n is negative.

Crude NSa. When dealing with the crude bandit algorithm (*i.e.*, when $\rho_1 = 0$, see (2.2.7)), the drift is reduced to φ_n^1 . One can check that $\varphi_n^1(y)$ is negative *iff*

$$\epsilon_n - \pi(1 - \gamma_n y) < 0 \iff y < \gamma_n^{-1} - \frac{\epsilon_n}{\pi\gamma_n} \iff x > \frac{\epsilon_n}{\pi}$$

where $x = 1 - \gamma_n y$. This means that when x is close to 0 (in some sense depending on n , π and γ_1), φ_n^1 becomes positive and Y_n has a tendency to increase. In others words, the dynamical system $(Y_n)_{n \geq 1}$ has no *mean-reverting* when Y_n is far from 0. The fact that the crude bandit algorithm does not always converge to the good target can be understood as a consequence of this remark.

Penalized and Over-Penalized NSa. When the drift φ_n contains a non zero penalty, the second term $-\varphi_n^2$ may help the dynamics to not be repulsive when x is close to 0, *i.e.* when y is larger than $1/\gamma_n$. It can be checked that $\kappa_\sigma(0) = 1 - \sigma p_2$ and :

$$\lim_{n \rightarrow +\infty} \varphi_n(\gamma_n^{-1}) = \frac{1}{2\gamma_1^2} - \frac{\gamma_1}{\rho_1}(1 - \sigma p_2).$$

This quantity is negative under the condition :

$$1 - \sigma p_2 > \frac{\rho_1}{2\gamma_1^3}. \quad (2.4.24)$$

But, in order to obtain a uniform bound on the regret, this constraint must be satisfied independently of p_2 . When $\sigma = 1$, *i.e.* in the standardly penalized case, one remarks that for any choice of ρ_1 and γ_1 , this is only possible if $\rho_1/(2\gamma_1^3) > 1 - p_2$. At this stage, one can thus understand the over-penalization as a way of controlling uniformly (in p_2) the negativity of φ_n far from $y = 0$ (see Figure 2.4).

In view of the main results, there are still two problems. The first one is that even in the over-penalized case, Inequality (2.4.24) implies some constraints on γ_1 and ρ_1 , which do not appear in

Theorem 2.3.2. The second one which is more embarrassing for the study of $(\mathbb{E}[Y_n])_{n \geq 1}$ is that, near $y = 0$, φ_n is positive since $\varphi_n(0) = 1 - \sigma p_1$ (see Figure 2.4). This repulsive behavior near $y = 0$ can be understood as the counterpart induced by the penalization. In order to bypass the two previous problems, the main argument will be the *increase of exponent* (see next section) where we show that we can replace the study of $(\mathbb{E}[Y_n])_{n \geq 1}$ by the one of a sequence which both has a nicer behavior near $y = 0$ and alleviates the constraint (2.4.24).

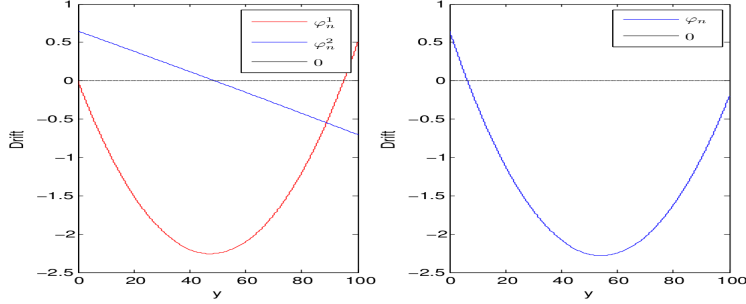


FIGURE 2.4 – Drift decomposition (left) and global (right) when $y \in [0, \frac{1}{\gamma_n}]$ with $\gamma_1 = \rho_1 = 1$, $p_1 = 0.7$, $p_2 = 0.6$, $\sigma = 0.5$.

2.4.3 Increase of exponent

We introduce the sequence $(Z_n^{(r)})_{n \geq 0}$ defined by :

$$\forall n \geq 1 \quad Z_n^{(r)} = \frac{(1 - X_n)^r}{\gamma_n}. \quad (2.4.25)$$

At this stage, one can first remark that *a.s.*, for every $r \geq 1$, $Z_n^{(r)} \leq Z_n^{(1)} = Y_n$. One can thus guess that the difficulties tackled at the end of the previous section will be easier to overcome for $(\mathbb{E}[Z_n^{(r)}])_{n \geq 1}$ with $r > 1$. Of course, this remark has an interest if conversely, one is able to relate the control of $\mathbb{E}[Y_n]$ to those of $\mathbb{E}[Z_n^{(r)}]$, $r \geq 1$.

This is the purpose of Proposition 2.4.1 where taking advantage of the structure of the algorithm, one shows that for every $r \geq 1$, $\mathbb{E}[Z_n^{(r)}]$ can be controlled by a function of $\mathbb{E}[Z_n^{(r+1)}]$.

Let us define the bounded function h_r on $[0, 1]$:

$$\forall \gamma \in [0, 1] \quad h_r(\gamma) = \frac{(1 + \gamma)^r - 1 - r\gamma}{r\gamma^2}. \quad (2.4.26)$$

Proposition 2.4.1. Let $r \in \mathbb{N}^*$, $\gamma_1 \in (0, 1)$ and $0 < \epsilon \leq \epsilon_0 = \frac{1}{3}$, and set

$$n_0(\epsilon, \pi, \gamma_1) := \left\lceil \frac{1}{4\epsilon^2 \gamma_1^2 \pi^2} \right\rceil + 1. \quad (2.4.27)$$

Then, if $2\epsilon\gamma_1^2(r - \epsilon) \leq 1$,

$$\sup_{n \geq n_0} \mathbb{E}Z_n^{(r)} \leq \mathbb{E}Z_{n_0}^{(r)} + \frac{r}{\pi(r - \epsilon)} \left[\tilde{\rho}_1 + h_r(\gamma_{n_0}) + \pi \sup_{n \geq n_0} \mathbb{E}[Z_n^{(r+1)}] \right].$$

In particular, for $r = 1, 2$, the previous inequality holds for every $\gamma_1 \in (0, 1)$ and $\epsilon \in (0, 1/3]$.

Remark 2.4.1. Note that the above result induces some constraints on ϵ and γ . These constraints, which allow us to manage the constants of the inequality, are mainly adapted to the proof of Theorem 2.3.2 (b). In fact, in the proofs of Theorems 2.3.1 and 2.3.2 (a), we will need to rewrite the above property in a slightly different way (see Section 2.4.5 for details).

Proof For any integer $r > 0$ and $n \geq 0$, the binomial formula applied to (2.4.22) leads to

$$\begin{aligned} (1 - X_{n+1})^r &= (1 - X_n - \Delta X_{n+1})^r \\ &= (1 - X_n)^r - r(1 - X_n)^{r-1} \Delta X_{n+1} \\ &\quad + \sum_{j=0}^{r-2} \binom{r}{j} (1 - X_n)^j (-\Delta X_{n+1})^{r-j}. \end{aligned}$$

where $\sum_{\emptyset} = 0$ and $\Delta X_{n+1} = X_{n+1} - X_n = \gamma_{n+1}[h(X_n) + \rho_{n+1}\kappa_{\sigma}(X_n) + \Delta M_{n+1}]$. From the definition of h given in (2.4.20), we get

$$(1 - x)^{r-1}[h(x) + \rho_{n+1}\kappa_{\sigma}(x)] = \pi x(1 - x)^r + \rho_{n+1}\kappa_{\sigma}(x)(1 - x)^{r-1}.$$

If we define now

$$\beta_n^{(r)} = -r\rho_{n+1}(1 - X_n)^{r-1}\kappa_{\sigma}(X_n) + \frac{1}{\gamma_{n+1}} \sum_{j=0}^{r-2} \binom{r}{j} (1 - X_n)^j (-\Delta X_{n+1})^{r-j}, \quad (2.4.28)$$

we can then conclude using (2.4.25) that

$$\begin{aligned} Z_{n+1}^{(r)} &= Z_n^{(r)} \frac{\gamma_n}{\gamma_{n+1}} - \gamma_n r \pi X_n Z_n^{(r)} + \beta_n^{(r)} - r(1 - X_n)^{r-1} \Delta M_{n+1} \\ &= Z_n^{(r)} \left(1 + \gamma_n \left[\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} - r\pi X_n \right] \right) + \beta_n^{(r)} - r(1 - X_n)^{r-1} \Delta M_{n+1} \\ &= Z_n^{(r)} (1 + \gamma_n [\epsilon_n - r\pi X_n]) + \beta_n^{(r)} - r(1 - X_n)^{r-1} \Delta M_{n+1} \\ &= Z_n^{(r)} (1 + \gamma_n [\epsilon_n - r\pi]) + r\pi\gamma_n(1 - X_n)Z_n^{(r)} + \beta_n^{(r)} - r(1 - X_n)^{r-1} \Delta M_{n+1} \\ &= Z_n^{(r)} (1 + \gamma_n [\epsilon_n - r\pi]) + r\pi\gamma_n Z_n^{(r+1)} + \beta_n^{(r)} - r(1 - X_n)^{r-1} \Delta M_{n+1}. \end{aligned} \quad (2.4.29)$$

The formulation above is important : it exhibits a contraction of $(1 + \gamma_n [\epsilon_n - r\pi])$ on $Z_n^{(r)}$ that can be used jointly with an upper bound of $Z_n^{(r+1)}$ and a simple majorization of $\beta_n^{(r)}$. In this view, we study (2.4.28) : $|\Delta X_{n+1}| \leq \gamma_{n+1}$ a.s. and (2.4.21) yields $|\kappa_{\sigma}(x)| \leq (1 - \sigma p_2)$. Now, with h_r given in (2.4.26), we get

$$\beta_n^{(r)} \leq r\tilde{\rho}_1\gamma_{n+1} + \sum_{j=0}^{r-2} \binom{r}{j} (\gamma_{n+1})^{r-j-1} \leq r(\tilde{\rho}_1 + h_r(\gamma_{n+1}))\gamma_{n+1}.$$

For any $\epsilon \in (0, 1)$, we can see in (2.4.29) that the contraction coefficient can be useful as soon as n is large enough. More precisely, using (2.4.23), we see that

$$\epsilon_n \leq \pi\epsilon \iff n \geq n_0(\epsilon, \pi, \gamma_1) := \left\lfloor \frac{1}{4\epsilon^2\gamma_1^2\pi^2} \right\rfloor + 1.$$

Then, for every $n \geq n_0(\epsilon, \pi, \gamma_1)$,

$$1 + \gamma_n [\epsilon_n - r\pi] \leq 1 - \alpha_r\gamma_n \quad \text{with} \quad \alpha_r = \pi(r - \epsilon).$$

In the sequel, we will omit the dependence of n_0 in $(\epsilon, \pi, \gamma_1)$ and will just use the notation n_0 . Also remark that under the condition $2\epsilon\gamma_1^2(r - \epsilon) \leq 1$, we have $\alpha_r\gamma_j < 1$ for every $\pi \in (0, 1)$ and for every $j \geq n_0$ (one can in particular check that $2\epsilon\gamma_1^2(r - \epsilon) \leq 1$ is true for every $\epsilon \in (0, 1/3)$ and $\gamma_1 \in (0, 1)$ if $r = 1, 2$). Thus, by a simple recursion based on (2.4.29), one obtains for every $n \geq n_0 + 1$,

$$\mathbb{E}(Z_n^{(r)}) \leq \mathbb{E}(Z_{n_0}^{(r)}) \prod_{j=n_0}^{n-1} (1 - \alpha_r\gamma_j) + \sum_{j=n_0}^{n-1} \left(r\pi\gamma_j \mathbb{E}(Z_j^{(r+1)}) + \beta_j^{(r)} \right) \prod_{l=j}^{n-1} (1 - \alpha_r\gamma_l)$$

If we call $\Theta_r = r \left(\pi \sup_{j \geq n_0} \left(\mathbb{E}(Z_j^{(r+1)}) + \tilde{\rho}_1 + h_r(\gamma_j) \right) \right)$, an iteration of the previous inequality yields :

$$\mathbb{E}(Z_n^{(r)}) \leq \mathbb{E}(Z_{n_0}^{(r)}) + \Theta_r \sum_{j=n_0}^{n-1} \gamma_j \prod_{l=j}^{n-1} (1 - \alpha_r\gamma_l).$$

We aim to apply Lemma 2.7.1 (deferred to the appendix section) to the last term. It is possible as soon as

$$n_0 \geq \frac{1}{(\alpha_r\gamma_1)^2}.$$

This last condition is fulfilled for any $r \geq 1$ when $\frac{1}{4\epsilon^2\gamma_1^2\pi^2} \geq \frac{1}{(1-\epsilon)^2\pi^2\gamma_1^2}$, *i.e.* when $\epsilon \leq 1/3$.

Then, by Lemma 2.7.1, one deduces that $\forall \epsilon \leq 1/3$ and $\forall n \geq n_0$:

$$\sup_{n \geq n_0} \mathbb{E}Z_n^{(r)} \leq \mathbb{E}Z_{n_0}^{(r)} + \frac{r}{\pi(r - \epsilon)} \left[\tilde{\rho}_1 + h_r(\gamma_{n_0}) + \pi \sup_{n \geq n_0} Z_n^{(r+1)} \right].$$

□

On the basis of the last proposition and a recursive argument, we can now deduce the following key observations.

Corollary 2.4.1. Assume that $\epsilon \in (0, 1/3)$, $\gamma_1 \in (0, 1)$ and that n_0 is defined in (2.4.27). Then,

$$\begin{aligned} \sup_{n \geq n_0} \mathbb{E}[Y_n] &\leq \mathbb{E}[Z_{n_0}^{(1)}] + \frac{\mathbb{E}[Z_{n_0}^{(2)}]}{1 - \epsilon} + \frac{1}{\pi(1 - \epsilon)} \left[\tilde{\rho}_1 + \frac{\tilde{\rho}_1}{1 - \epsilon/2} + \frac{1}{2(1 - \epsilon/2)} \right] \\ &\quad + \frac{\sup_{n \geq n_0} \mathbb{E}Z_n^{(3)}}{(1 - \epsilon)(1 - \epsilon/2)}. \end{aligned} \quad (2.4.30)$$

Remark 2.4.2. As in Proposition 2.4.1, this property is mainly written in view of Theorem 2.3.2 (b) where we only need to use the increase of exponent for $r = 1, 2$. For Theorems 2.3.1 and 2.3.2 (a) with $\sigma \in (0, 1)$, we will need to use it for large values of r .

2.4.4 Bound for $(\mathbb{E}(Z_n^{(3)}))_{n \geq n_0}$

As seen in Corollary 2.4.1, our next task is to bound $\mathbb{E}(Z_n^{(3)})$ for $n \geq n_0$ to obtain a tractable application of Equation (2.4.30). Such a bound is reached through careful inspection of the increments $\Delta Z_{n+1}^{(3)} := Z_{n+1}^{(3)} - Z_n^{(3)}$.

Lemma 2.4.1 (Decomposition of $Z_n^{(3)}$). For every $n \geq 1$,

$$\mathbb{E}[\Delta Z_{n+1}^{(3)} | \mathcal{F}_n] = \gamma_{n+1}(1 - X_n)P_n(X_n) + \Delta R_n,$$

where for every $n \in \mathbb{N}$, P_n is a polynomial function defined by

$$\begin{aligned} P_n(x) &= \frac{(1-x)^2}{\gamma_{n+1}} (\epsilon_n - 3\pi x) - 3\tilde{\rho}_1(1-x)\kappa_\sigma(x) + 3(x(1-x)^2 p_1 + x^2(1-x)p_2) \\ &\quad + \gamma_{n+1}(-x(1-x)^2 p_1 + x^3 p_2), \end{aligned} \quad (2.4.31)$$

and if $(\gamma_1, \rho_1) \in (0, 1]^2$, then

$$\forall n \geq n_0, \quad \Delta R_n \leq (1 - \sigma p_2) [3\gamma_{n+1}\rho_{n+1}^2 + \gamma_{n+1}^2\rho_{n+1}^3] \quad a.s.$$

Remark 2.4.3. — The keypoint is that $\gamma_k = \gamma_1 k^{-1/2}$ and, therefore, the series $\sum_{n \geq 1} \Delta R_k$ is uniformly bounded, regardless of the value of π . This will be enough to obtain a competitive upper bound of the regret. With the choice of n_0 given in (2.4.27), careful inspection of Lemma 2.4.1 leads to :

$$\sum_{k \geq n_0} \Delta R_k \leq 12\gamma_1^4 \tilde{\rho}_1^2 \epsilon \pi + \frac{16}{3} \gamma_1^8 \tilde{\rho}_1^3 \epsilon^3 \pi^3 \quad a.s. \quad (2.4.32)$$

— As in Remark 2.4.2, it should be noted that for Theorems 2.3.1 and 2.3.2 (a) with $\sigma \in (0, 1)$, we will need to use such a development with some larger values of r (see the end of this section for details).

Proof We again use Equation (2.4.29) and deduce that :

$$Z_{n+1}^{(3)} - Z_n^{(3)} = (1 - X_n)^3 (\epsilon_n - 3\pi X_n) - 3\tilde{\rho}_1 \gamma_{n+1} (1 - X_n)^2 \kappa_\sigma(X_n) \quad (2.4.33)$$

$$+ \frac{1}{\gamma_{n+1}} \sum_{j=0}^1 \binom{3}{j} (1 - X_n)^j (-\Delta X_{n+1})^{3-j} - 3(1 - X_n)^2 \Delta M_{n+1} \quad (2.4.34)$$

First, note that terms in Equation (2.4.33) are associated with the first two terms in the definition of P_n introduced in (2.4.31) up to a multiplication by $(1 - X_n)\gamma_{n+1}$.

Second, we can easily compute the expectations involved in the sum of Equation (2.4.34) since the events are all disjointed. On the one hand, when $j = 1$ we have

$$\begin{aligned} \frac{1}{\gamma_{n+1}} \mathbb{E}[(-\Delta X_{n+1})^2 | \mathcal{F}_n] &= \gamma_{n+1} \sigma (p_1 X_n (1 - X_n)^2 + p_2 (1 - X_n) X_n^2) \\ &+ \gamma_{n+1} (1 - \sigma) (p_1 X_n (1 - X_n - \rho_{n+1} X_n)^2 + p_2 (1 - X_n) (X_n - \rho_{n+1} (1 - X_n))^2) \\ &+ \gamma_{n+1} \rho_{n+1}^2 [X_n^3 (1 - p_1) + (1 - X_n)^3 (1 - p_2)]. \end{aligned}$$

Further computations yield :

$$\begin{aligned} \frac{1}{\gamma_{n+1}} \mathbb{E}[(-\Delta X_{n+1})^2 | \mathcal{F}_n] &= \gamma_{n+1} X_n (\sigma p_1 (1 - X_n)^2 + \sigma p_2 X_n^2) + \Delta A_n^{(1)} \\ &+ \underbrace{\gamma_{n+1} \rho_{n+1}^2 [X_n^3 (1 - \sigma p_1) + (1 - X_n)^3 (1 - \sigma p_2)]}_{:= \Delta R_n^{(1)}} \end{aligned}$$

with $\Delta A_n^{(1)} = -2\rho_{n+1}\gamma_{n+1}X_n(1 - X_n)(1 - \sigma)(X_n p_1 + (1 - X_n)p_2)$. On the other hand, we can also compute the term when $j = 0$:

$$\begin{aligned} \frac{1}{\gamma_{n+1}} \mathbb{E}[(-\Delta X_{n+1})^3 | \mathcal{F}_n] &= \gamma_{n+1}^2 X_n (1 - X_n) (p_2 X_n^2 - p_1 (1 - X_n)^2) \\ &+ \Delta A_n^{(2)} + \underbrace{\gamma_{n+1}^2 \rho_{n+1}^3 [X_n^4 (1 - \sigma p_1) - (1 - X_n)^4 (1 - \sigma p_2)]}_{:= \Delta R_n^{(2)}} \end{aligned}$$

with $\Delta A_n^{(2)} \leq 3\gamma_{n+1}^2 \rho_{n+1} (1 - \sigma) X_n (1 - X_n)^2 (\pi X_n + \rho_{n+1} (1 - X_n) p_2)$. Set $\Delta R_n^{(3)} = 3(1 - X_n) \Delta A_n^{(1)} + \Delta A_n^{(2)}$ and $\Delta R_n := 3(1 - X_n) \Delta R_n^{(1)} + \Delta R_n^{(2)} + \Delta R_n^{(3)}$. Plugging the previous controls into (2.4.34) yields

$$\mathbb{E}[\Delta Z_{n+1}^{(3)} | \mathcal{F}_n] \leq \gamma_{n+1} (1 - X_n) P_n(X_n) + \Delta R_n. \quad (2.4.35)$$

Note that $\Delta R_n^{(1)}$ can be upper bounded as follows :

$$3(1 - X_n) \Delta R_n^{(1)} \leq 3\gamma_{n+1} \rho_{n+1}^2 (1 - \sigma p_2) \max_{0 \leq t \leq 1} \left[\frac{1 - \sigma p_1}{1 - \sigma p_2} t^3 (1 - t) + (1 - t)^4 \right].$$

Since $1 - \sigma p_1 \leq 1 - \sigma p_2$, a study of the function shows that $at^3(1 - t) + (1 - t)^4$ when $a \in (0, 1)$ reaches its maximal value for $t = 0$. This leads to :

$$3(1 - X_n) \Delta R_n^{(1)} \leq 3\gamma_{n+1} \rho_{n+1}^2 (1 - \sigma p_2).$$

For $\Delta R_n^{(2)}$, we have $\Delta R_n^{(2)} \leq \gamma_{n+1}^2 \rho_{n+1}^3 (1 - \sigma p_2) \max_{0 \leq t \leq 1} \left[\frac{1 - \sigma p_1}{1 - \sigma p_2} t^4 - (1 - t)^4 \right]$, which involves an increasing function of t . Thus, we have

$$\Delta R_n^{(2)} \leq \gamma_{n+1}^2 \rho_{n+1}^3 (1 - \sigma p_1) \leq \gamma_{n+1}^2 \rho_{n+1}^3 (1 - \sigma p_2).$$

Finally, if γ_1 and $\rho_1 \in (0, 1]$, then one can readily check that

$$\Delta A_n^{(2)} \leq -3(1 - X_n) \Delta A_n^{(1)} \implies \Delta R_n^{(3)} \leq 0.$$

The result follows according to Equation (2.4.35). □

In order to bound $\sup_{n \geq n_0} \mathbb{E}(Z_n^{(3)})$, we now have to precisely study the polynomial function P_n and exhibit a mean reverting effect on its dynamics.

Proposition 2.4.2. Let $\epsilon \in (0, \frac{1}{3})$, $\tilde{\rho}_1 \leq \frac{227}{232}$ and $\frac{1}{3\sqrt{2}(1-\sigma)\tilde{\rho}_1} \leq \gamma_1^2 \leq \frac{3}{2(1+\tilde{\rho}_1)}$. Then

- i) The polynomial P_n given by (2.4.31) is negative on $[0, 1 - \frac{2(1+\tilde{\rho}_1)}{\pi} \gamma_{n+1}]$.
- ii) $Z_n^{(3)}$ satisfies

$$\begin{aligned} \sup_{n \geq n_0} \mathbb{E} Z_n^{(3)} &\leq \mathbb{E} Z_{n_0}^{(3)} + 12\gamma_1^4 \tilde{\rho}_1^2 \epsilon \pi + \frac{16}{3} \gamma_1^8 \tilde{\rho}_1^3 \epsilon^3 \pi^3 \\ &\quad + \frac{8\gamma_1^4 \epsilon (1 + \tilde{\rho}_1) [1 + (1 + \tilde{\rho}_1)[2 + 6\tilde{\rho}_1 + 12\gamma_1^2 \epsilon]]}{\pi}. \end{aligned}$$

Remark 2.4.4. The above result is given under some technical conditions that will lead to a sharp explicit bound. Nevertheless, the reader has to keep in mind that in view of the condition on σ , the “universal” bound on $(\mathbb{E}(Z_n^{(3)}))_{n \geq n_0}$ is only accessible when $\sigma < 1$, *i.e.* in the over-penalized case. When $\sigma = 1$, some bounds will be attainable only if p_2 is not too large (see (2.4.24) for a similar statement when $r = 1$), and in order to alleviate the constraint on p_2 , it will be necessary to take a larger exponent than $r = 3$ (see Subsection 2.4.5 for details).

Proof We first provide the proof of i). The function P_n introduced in (2.4.31) is a third degree polynomial and for $n \geq n_0$:

$$\begin{aligned} P_n(0) &= \frac{\epsilon_n}{\gamma_{n+1}} - 3\tilde{\rho}_1 \kappa_\sigma(0) \\ &\leq \frac{\gamma_n}{2\gamma_1^2 \gamma_{n+1}} - 3\tilde{\rho}_1 (1 - \sigma p_2) \\ &\leq \frac{\sqrt{1 + n_0^{-1}}}{2\gamma_1^2} - 3\tilde{\rho}_1 (1 - \sigma p_2) \end{aligned}$$

Since $p_2 < 1$, this last quantity is negative if :

$$\rho_1 \gamma_1 \geq \frac{1}{3\sqrt{2}(1 - \sigma)}. \quad (2.4.36)$$

In a same way, we can check that $P_n(1) = \gamma_{n+1} p_2 > 0$ and, therefore, P_n has one root in the interval

$(0, 1)$. Careful inspection of the leading coefficient (designated $a_n x^3$) of P_n in (2.4.31) shows that :

$$a_n = \left[3(1 + \sigma \tilde{\rho}_1) - \frac{3}{\gamma_{n+1}} - \gamma_{n+1} \right] \pi.$$

The leading coefficient a_n is negative as soon as $3(1 + \sigma \tilde{\rho}_1) \leq \frac{3}{\gamma_{n+1}}$. Again, the choice of n_0 in (2.4.27) shows that this last condition is fulfilled as soon as

$$\frac{1}{\epsilon} \geq 2\gamma_1 \pi (\gamma_1 + \sigma \rho_1). \quad (2.4.37)$$

It should however be noted that we have assumed $\epsilon \in (0, 1/3]$ so that $\frac{1}{\epsilon} \geq 3$. As a consequence, (2.4.36) and (2.4.37) are satisfied as soon as $(\gamma_1, \tilde{\rho}_1)$ satisfies

$$\frac{1}{3\sqrt{2}(1 - \sigma)\tilde{\rho}_1} \leq \gamma_1^2 \leq \frac{3}{2(1 + \tilde{\rho}_1)}$$

Hence, if (2.4.36) and (2.4.37) hold, P_n possesses one root in $(-\infty, 0)$ and another one in $(1, +\infty)$. Consequently, P_n has a unique root in $(0, 1)$. We now consider :

$$\xi_n = \frac{2(1 + \tilde{\rho}_1)}{\pi} \gamma_{n+1} := \xi \gamma_{n+1}.$$

We compute that :

$$\begin{aligned} & P_n(1 - \xi_n) \\ &= \frac{\xi_n^2}{\gamma_{n+1}} (\epsilon_n - 3\pi(1 - \xi_n)) - 3\tilde{\rho}_1 \xi_n [(1 - \sigma p_2) \xi_n^2 - (1 - \sigma p_1)(1 - \xi_n)^2] \\ & \quad + 3 [(1 - \xi_n) \xi_n^2 p_1 + (1 - \xi_n)^2 \xi_n p_2] + \gamma_{n+1} [(1 - \xi_n)^3 p_2 - \xi_n^2 (1 - \xi_n)]. \end{aligned}$$

Hence, replacing ξ_n by $\xi \gamma_{n+1}$ and simplifying by γ_{n+1} , we see that $P_n(1 - \xi_n)$ is negative when

$$\begin{aligned} & \overbrace{\frac{\xi^2 \epsilon_n}{(1 - \xi_n)} + 3\tilde{\rho}_1(1 - \sigma p_1)(1 - \xi_n)\xi + 3p_1 \gamma_{n+1} \xi^2 + 3p_2(1 - \xi_n)\xi + p_2(1 - \xi_n)^2}^{:= A_n(\xi)} \\ & \leq 3\pi \xi^2 + \underbrace{\frac{3\tilde{\rho}_1 \gamma_{n+1}^2 \xi^3 (1 - \sigma p_2)}{1 - \xi_n} + \gamma_{n+1}^2 \xi^2}_{:= B_n(\xi)}. \end{aligned}$$

From (2.4.23), we know that $\epsilon_n \leq \frac{\gamma_{n+1}}{2\gamma_1^2}$, and $1 - \xi_n \leq 1$ thus

$$A_n(\xi) \leq \xi^2 \gamma_{n+1} \left(\frac{1}{2\gamma_1^2(1 - \xi_n)} + 3p_1 \right) + 3\xi (\tilde{\rho}_1 + 1) + 1$$

In the meantime, we will use the simple lower bound $B_n(\xi) \geq 3\pi \xi^2$. We can check that $1 - \xi_n =$

$1 - \frac{2(1+\tilde{\rho}_1)\gamma_{n+1}}{\pi} \geq 1 - 4\epsilon(1+\tilde{\rho}_1)\gamma_1^2$ since $\gamma_{n_0} \leq 2\epsilon\gamma_1^2\pi$. Thus

$$\begin{aligned} A_n \left(\frac{2(1+\tilde{\rho}_1)}{\pi} \right) &\leq \frac{4(1+\tilde{\rho}_1)^2}{\pi^2} \gamma_{n+1} \left[3p_1 + \frac{1}{2\gamma_1^2 [1 - 4\epsilon(1+\tilde{\rho}_1)\gamma_1^2]} \right] + \frac{6(1+\tilde{\rho}_1)^2}{\pi} + 1 \\ &\leq \frac{(1+\tilde{\rho}_1)^2}{\pi} \left[24\epsilon\gamma_1^2 p_1 + \frac{4\epsilon}{1 - 4\epsilon(1+\tilde{\rho}_1)\gamma_1^2} + 7 \right] \end{aligned}$$

and

$$B_n \left(\frac{2(1+\tilde{\rho}_1)}{\pi} \right) \geq \frac{12(1+\tilde{\rho}_1)^2}{\pi}.$$

As a consequence, $P_n(1 - \xi_n)$ is negative if we have

$$5 \geq 24\epsilon\gamma_1^2 + \frac{4\epsilon}{1 - 4\epsilon(1+\tilde{\rho}_1)\gamma_1^2}$$

From the constraint on γ_1 , another computation shows that the above condition is fulfilled when $\epsilon^2 \frac{128(1+\tilde{\rho}_1)}{3} - \epsilon[84 + 40(1+\tilde{\rho}_1)] + 45 \geq 0$. We then observe that all values of ϵ in $(0, \frac{1}{3}]$ can be conveniently used when $\tilde{\rho}_1 \leq \frac{227}{232}$.

To obtain *ii*), the main idea is to use the sharp estimation of the sign of P_n on $[0, 1]$ and to obtain an upper bound of $\mathbb{E}Z_n^{(3)}$. Note that :

$$\begin{aligned} &\sup_{0 \leq t \leq 1} \gamma_{n+1}(1-t)P_n(t) \\ &= \gamma_{n+1} \sup_{1-\xi_n \leq t \leq 1} (1-t)P_n(t) \\ &= \gamma_{n+1} \sup_{1-\xi_n \leq t \leq 1} \left\{ (1-t)^3 [\epsilon_n - 3\pi t] - 3\tilde{\rho}_1(1-t)^2 \kappa_\sigma(t) \right. \\ &\quad \left. + 3[t(1-t)^3 p_1 + t^2(1-t)^2 p_2] + \gamma_{n+1} [-t(1-t)^3 p_1 + t^3(1-t)p_2] \right\} \end{aligned}$$

We have seen in the proof of *i*) that $t \in [1 - \xi_n, 1] \implies \epsilon_n \leq 3\pi t$. Hence, using $\kappa_\sigma(t) \geq -(1 - \sigma p_1)t^2$, we have

$$\begin{aligned} &\sup_{0 \leq t \leq 1} \gamma_{n+1}(1-t)P_n(t) \\ &\leq \gamma_{n+1} [3\tilde{\rho}_1(1 - \sigma p_1)\xi_n^2 + 3p_1\xi_n^3 + p_2\xi_n^2 + \gamma_{n+1}\xi_n] \\ &\leq \frac{C_1(\tilde{\rho}_1, p_1, p_2, \sigma)}{\pi^2} \gamma_{n+1}^3 + \frac{C_2(\tilde{\rho}_1, p_1)}{\pi^3} \gamma_{n+1}^4 \end{aligned}$$

with $C_1(\tilde{\rho}_1, p_1, p_2, \sigma) = (1+\tilde{\rho}_1)(12\tilde{\rho}_1(1+\tilde{\rho}_1)(1-\sigma p_1) + 4p_2(1+\tilde{\rho}_1) + 2\pi)$ and $C_2(\tilde{\rho}_1, p_1) = 24p_1(1+$

$\tilde{\rho}_1)^3$ shortened in C_1 and C_2 below. We apply Lemma 2.4.1 to upper bound $\sup_{n \geq n_0} \mathbb{E}Z_n^{(3)}$:

$$\begin{aligned}
& \sup_{n \geq n_0} \mathbb{E}Z_n^{(3)} \\
& \leq \mathbb{E}Z_{n_0}^{(3)} + \sup_{n \geq n_0} \mathbb{E} \sum_{k=n_0}^n \Delta Z_{n+1}^{(3)} \\
& \leq \mathbb{E}Z_{n_0}^{(3)} + \sup_{n \geq n_0} \mathbb{E} \left[\sum_{k=n_0}^n \gamma_{k+1}(1 - X_k)P_k(X_k) + \Delta R_k \right] \\
& \leq \mathbb{E}Z_{n_0}^{(3)} + \frac{C_1}{\pi^2} \sum_{k=n_0}^{\infty} \gamma_{k+1}^3 + \frac{C_2}{\pi^3} \sum_{k=n_0}^{\infty} \gamma_{k+1}^4 + \sum_{k=n_0}^{\infty} \mathbb{E}[\max(\Delta R_k, 0)].
\end{aligned}$$

Using a simple comparison argument with the integrals $\int_{n_0}^{\infty} t^{-\alpha} dt$, we obtain :

$$\sum_{k=n_0}^{\infty} \gamma_{k+1}^3 \leq 2\gamma_1^3 n_0^{-1/2} \leq 4\gamma_1^4 \epsilon \pi \quad \text{and} \quad \sum_{k=n_0}^{\infty} \gamma_{k+1}^4 \leq \gamma_1^4 n_0^{-1} \leq 4\gamma_1^6 \epsilon^2 \pi^2.$$

We then deduce that :

$$\sup_{n \geq n_0} \mathbb{E}Z_n^{(3)} \leq \mathbb{E}Z_{n_0}^{(3)} + \frac{4\gamma_1^4 \epsilon C_1}{\pi} + \frac{4\gamma_1^6 \epsilon^2 C_2}{\pi} + \sum_{k=n_0}^{\infty} \mathbb{E}[\max(\Delta R_k, 0)].$$

The result now follows using (2.4.32).

□

Explicit bound. We can now conclude the proof of Theorem 2.3.2.

Proof of Theorem 2.3.2 (b). We consider the extreme over-penalized case obtained with $\sigma = 0$ and use a power increment until $r = 3$. Recall that $n_0 := n_0(\epsilon, \pi, \gamma_1)$ is defined by (2.4.27). In particular, $\sqrt{n_0 - 1} \leq (2\epsilon\gamma_1\pi)^{-1}$ and for $i = 1, 2, 3$, $\pi\mathbb{E}[Z_{n_0}^{(i)}] \leq (2\epsilon\gamma_1^2)^{-1} + (\gamma_1)^{-1}$. Taking the results of Proposition 2.4.2 ii) and Corollary 2.4.1 and plugging them into (2.4.19), a series of computations yields :

$$\frac{\sup_{p_1 \geq p_2} \bar{R}_n}{\sqrt{n}} \leq c(\gamma_1, \tilde{\rho}_1, \epsilon) := T_1(\gamma_1, \tilde{\rho}_1, \epsilon) + \frac{2\gamma_1}{(1 - \epsilon)(1 - \epsilon/2)} T_2(\gamma_1, \tilde{\rho}_1, \epsilon), \quad (2.4.38)$$

where

$$\begin{aligned}
T_1(\gamma_1, \tilde{\rho}_1, \epsilon) &= \frac{1}{2\epsilon\gamma_1} + \left(\frac{1}{\epsilon\gamma_1} + 2 \right) \left(1 + \frac{1}{1 - \epsilon} + \frac{1}{(1 - \epsilon)(1 - \epsilon/2)} \right) \\
&\quad + 2\rho_1 \left(\frac{1}{1 - \epsilon} + \frac{1}{(1 - \epsilon)(1 - \epsilon/2)} \right) + \frac{\gamma_1}{(1 - \epsilon)(1 - \epsilon/2)},
\end{aligned}$$

and

$$\begin{aligned} T_2(\gamma_1, \tilde{\rho}_1, \epsilon) \\ = \gamma_1^4 \left[8\epsilon(1 + \tilde{\rho}_1) (1 + (1 + \tilde{\rho}_1)(2 + 6\tilde{\rho}_1 + 12\gamma_1^2\epsilon)) + 12\tilde{\rho}_1^2\epsilon + \frac{16}{3}\gamma_1^4\tilde{\rho}_1^3\epsilon^3 \right]. \end{aligned}$$

Theorem 2.3.2(b) follows by minimizing $(\gamma_1, \tilde{\rho}_1, \epsilon) \mapsto c(\gamma_1, \tilde{\rho}_1, \epsilon)$ under the constraints :

$$\epsilon \leq 1/3, \frac{1}{3\sqrt{2}\tilde{\rho}_1} \leq \gamma_1^2 \leq \frac{3}{2(1 + \tilde{\rho}_1)}, \tilde{\rho}_1 \leq 227/232.$$

The “best” upper bound was obtained by setting $\gamma_1 = 0.89, \tilde{\rho}_1 = 0.38, \epsilon = 1/9$, leading to the regret upper bound

$$\bar{R}_n \leq 31.1\sqrt{2}\sqrt{n} \leq 44\sqrt{n}.$$

□

2.4.5 Proof of Theorems 2.3.1 and 2.3.2 (a)

We prove these results together. We thus consider $\gamma_1 \in (0, 1)$, $\rho_1 \in (0, 1)$ and $\sigma \in [0, 1]$. A variant of Proposition 2.4.1 concerning the increase of exponent is still valid. First, it can be observed that if we set $\epsilon_r = r - 1/2$ (so that $\alpha_r = \pi/2$), then Lemma 2.7.1 can be applied with $\tilde{n} \geq (\frac{\pi}{2}\gamma_1)^{-2}$. Thus, we set $n_0(\lambda) := \lfloor \frac{\lambda^2}{\pi^2} \rfloor + 1$ with $\lambda \geq 2\gamma_1^{-1}$. After a simple adaptation of the proof of Proposition 2.4.1, it can be deduced that for every $r \geq 1$,

$$\sup_{n \geq n_0(\lambda)} \mathbb{E}Z_n^{(r)} \leq \mathbb{E}Z_{n_0}^{(r)} + \frac{2r}{\pi} \left[\tilde{\rho}_1 + h_r(\gamma_{n_0(\lambda)}) + \pi \sup_{n \geq n_0(\lambda)} Z_n^{(r+1)} \right].$$

By an iteration, it follows by using the fact that $\pi \mathbb{E}[Z_{n_0(\lambda)}^{(i)}] \leq \pi \gamma_{n_0(\lambda)}^{-1} \leq \gamma_1^{-1}(\lambda + 1)$ that for every $r \geq 1$, some constants $C_r^1(\lambda)$ and $C_r^2(\lambda)$ exist (depending only on σ, γ_1 and ρ_1) such that,

$$\sup_{n \geq n_0(\lambda)} \pi \mathbb{E}[Y_n] \leq C_r^1(\lambda) + C_r^2(\lambda) \pi \sup_{n \geq n_0(\lambda)} \mathbb{E}Z_n^{(r+1)}. \quad (2.4.39)$$

It remains to upper bound $\sup_{n \geq n_0(\lambda)} \mathbb{E}Z_n^{(r)}$ for r large enough. Once again, a simple adaptation of the proof of Lemma 2.4.1 for $r \geq 3$ yields :

$$\mathbb{E}[\Delta Z_{n+1}^{(r)} | \mathcal{F}_n] = \gamma_{n+1}(1 - X_n)^{r-1} P_n^{(r)}(X_n) + \Delta R_n^{(r)}.$$

with

$$\begin{aligned} P_n^{(r)}(x) \\ = \frac{(1-x)^2}{\gamma_{n+1}} (\epsilon_n - r\pi x) - r\tilde{\rho}_1(1-x)\kappa_\sigma(x) + \binom{r}{r-2} (x(1-x)^2 p_1 + x^2(1-x)p_2) \\ + \gamma_{n+1} \binom{r}{r-3} (-x(1-x)^2 p_1 + x^3 p_2) \end{aligned} \quad (2.4.40)$$

and $\Delta R_n^{(r)} \leq C_r \gamma_{n+1}^3$ (where C_r does not depend on π). We want to prove that $P_n^{(r)}$ is negative on $[0, 1 - \xi_n]$ with $\xi_n = \xi \gamma_{n+1} \in (0, 1)$ where ξ is a constant to be calibrated. We follow the lines of the proof of Proposition 2.4.2, but we can use some rougher arguments since we are not looking for explicit constants. First, $P_n^{(r)}(0) = \frac{\epsilon_n}{\gamma_{n+1}} - r \tilde{\rho}_1 \kappa_\sigma(0)$, so that :

$$P_n^{(r)}(0) < 0 \iff \gamma_1 \rho_1 > \frac{1}{r \sqrt{2}(1 - \sigma p_2)}.$$

On the one hand, for every $\sigma < 1$, it is possible to find an r sufficiently large for which this condition holds. On the other hand, when $\sigma = 1$ (case of Theorem 2.3.1), we then need to assume that a $\delta > 0$ exists such that $p_2 < 1 - \delta$ (in this case, the condition is satisfied if $r > (\gamma_1 \rho_1 \sqrt{2} \delta)^{-1}$). For such an r , it can be observed that the leading coefficient $a_n^{(r)}$ (related to x^3) is :

$$a_n^{(r)} = \left(-\frac{r}{\gamma_{n+1}} + \binom{r}{r-2} + r \sigma \tilde{\rho}_1 - \gamma_{n+1} \right) \pi.$$

It can therefore be deduced that $a_n^{(r)}$ is negative for every $n \geq n_1^\sigma$ where :

$$n_1^\sigma := \left\lceil \gamma_1^2 \left(\frac{r-1}{2} + \sigma \tilde{\rho}_1 \right)^2 \right\rceil.$$

Assume that $\lambda \geq \sqrt{n_1^\sigma}$ in order to obtain $n_0(\lambda) \geq n_1^\sigma$. Since $P_n^{(r)}(1) = \gamma_{n+1} \binom{r}{r-3} p_2 > 0$ and $\deg(P_n^{(r)}) = 3$, it follows that $P_n^{(r)}$ has exactly one root in $(0, 1)$ for every $n \geq n_0$ and that $P_n^{(r)}$ is negative on $[0, 1 - \xi_n]$ as soon as $P_n^{(r)}(1 - \xi_n) < 0$. Let n be such that $\xi \gamma_{n+1} \leq 1/2$. Then, some rough estimations yield that $P_n^{(r)}(1 - \xi_n)$ is negative if

$$\frac{r\pi}{2} \xi^2 - c_r \xi - 1 > 0,$$

where c_r is a constant that does not depend on π . We then check that another constant η_r exists such that the previous property is fulfilled if $\xi \geq \eta_r/\pi$. Then, $P_n^{(r)}(1 - \frac{\eta_r}{\pi} \gamma_{n+1}) < 0$ is negative as soon as $\xi \gamma_{n+1} < 1/2$. This is true for every $n \geq n_0(\lambda)$ as soon as $\lambda \geq 2\gamma_1 \eta_r$. We can conclude from what precedes that an $r \geq 3$ and $\lambda > 0$ exist such that for every $n \geq n_0(\lambda)$, for every $(p_1, p_2) \in [0, 1]^2$, such that $p_1 > p_2$ (resp. $p_1 > p_2$ and $p_2 < 1 - \delta$) if $\sigma < 1$ (resp. if $\sigma = 1$)

$$\mathbb{E}[\Delta Z_{n+1}^{(r)}] \leq \gamma_{n+1} \sup_{t \in [1 - \frac{\eta_r}{\pi} \gamma_{n+1}, 1]} (1 - t) P_n^{(r)}(t) + C_r \gamma_{n+1}^3.$$

Using $\gamma_{n+1} \leq \pi/\lambda$ if $n \geq n_0(\lambda)$, a constant C_λ exists such that on

$$\forall t \in [1 - \frac{\eta_r}{\pi} \gamma_{n+1}, 1] \quad P_n^{(r)}(t) \leq C_\lambda \gamma_{n+1} / \pi.$$

Under the previous conditions, we deduce

$$\sup_{\pi} \left(\pi \sup_{n \geq n_0(\lambda)} \mathbb{E} Z_n^{(r)} \right) \leq \sup_{\pi} \left(\pi \sum_{n \geq n_0} C \gamma_{n+1}^3 (\pi^{-2} + \pi^{-1}) \right) < +\infty.$$

The result follows by plugging this inequality into (2.4.39).

2.5 Almost sure and weak limit of the over-penalized bandit

We provide here the proofs of Propositions 2.3.1 and 2.3.2. For the sake of simplicity, we restrict our study to $\sigma = 1$ (always over-penalization of the bandit), and the argument can be adapted for any values of $\sigma \in (0, 1]$.

2.5.1 A.s. convergence of the multi-armed bandit (Proposition 2.3.1)

Recall first that $X_n = (X_n^1, \dots, X_n^d)$, the multi-armed penalized bandit (2.2.13) makes it possible to define for $i \in \{2, \dots, d\}$,

$$X_{n+1}^i = X_n^i + \gamma_{n+1} h_i(X_n) + \gamma_{n+1} \rho_{n+1} \kappa_i(X_n) + \gamma_{n+1} \Delta M_{n+1,i},$$

where the main part of the drift h_i is defined as

$$h_i(x_1, \dots, x_d) = (1 - x_i) x_i p_i - x_i \sum_{j \neq i} x_j p_j,$$

and the penalty drift is

$$\kappa_i(x_1, \dots, x_d) = -x_i^2(1 - p_i) + \frac{1}{d-1} \sum_{j \neq i} x_j^2(1 - p_j).$$

Hence, the martingale increment is simply obtained as

$$\begin{aligned} \Delta M_{n+1,i} &= ((1 - X_n^i) 1_{I_{n+1}=i, A_{n+1}^i} - X_n^i \sum_{j \neq i} 1_{I_{n+1}=j, A_{n+1}^j} - h_i(X_n)) \\ &- \rho_{n+1} (X_n^i 1_{I_{n+1}=i, (A_{n+1}^i)^c} - \frac{1}{d-1} \sum_{j \neq i} X_n^j 1_{I_{n+1}=j, (A_{n+1}^j)^c} + \kappa_i(X_n)) \end{aligned}$$

Proof of Proposition 2.3.1. We start by (i) and identify the stationary points of the ODE method. The ODE $\dot{x} = h(x)$ possesses a finite number of equilibria that can be easily identified. We begin by solving the equation $h_1(x) = 0$. Since

$$h_1(x) = x_1 \sum_{j=2}^d x_j (p_1 - p_j) \geq 0,$$

we either have $x_1 = 1$ and $x_2 = \dots = x_d = 0$ or $x_1 = 0$.

Then, the equation $h_2(x) = 0$ with $x_1 = 0$ may be reduced to

$$x_2 \sum_{j=3}^d x_j (p_2 - p_j) \geq 0.$$

The same argument leads to $x_2 = 1$ or $x_2 = 0$ and a straightforward recursion shows that the

equilibria of the ODE are $(\delta^i)_{1 \leq i \leq d}$, with $(\delta^i)_{1 \leq i \leq d}$ defined as

$$\delta_i^i = 1 \quad \text{and} \quad \delta_j^i = 0 \quad \forall j \neq i.$$

Let us emphasize that to discriminate among these equilibria, it is not possible to use the second derivative criterion that relies on $\left(\frac{\partial h_i}{\partial x_j}\right)_{i,j}$ to establish their stability. Instead, it is possible to check that δ^1 fulfills the Lyapunov certificate with the function $V(x) = \frac{1}{2}(x_2^2 + \dots + x_d^2)$. If we denote $h = (h_1, \dots, h_d)$, we then have :

$$\langle \nabla V(x), h(x) \rangle = \sum_{j=2}^d x_j^2 \sum_{k \neq j} x_k (p_j - p_k).$$

Considering x in a closed neighborhood of δ^1 defined as $x_j \leq \epsilon/d$, $\forall j \geq 2$ (implying that $x_1 > 1 - \epsilon$), we see that :

$$\begin{aligned} \langle \nabla V(x), h(x) \rangle &= x_1 \sum_{j=2}^d x_j^2 (p_j - p_1) + \sum_{k=2}^d x_k^2 \sum_{j \neq k, j > 1} x_j (p_k - p_j) \\ &\leq -(1 - \epsilon)(p_1 - p_2) \sum_{j=2}^d x_j^2 + \epsilon \sum_{j=2}^d x_j^2, \end{aligned}$$

and the term above is negative as soon as ϵ is chosen such that :

$$\epsilon < \frac{p_1 - p_2}{p_1 - p_2 + 1}.$$

In contrast, the other equilibria $(\delta^j), j \neq 1$ are unstable : this can be easily deduced from the unstability of the two-armed bandit by testing the first arm vs. the arm j .

Since the martingale increment $\Delta M_{n+1,i}$ is uniformly bounded, we can apply the Kushner-Clark theorem (see [72]) and can conclude that $(X_n^i)_{n \geq 0}$ either converges to 1 or 0 a.s. As a consequence, it is also true that $(X_n)_{n \geq 0}$ converges a.s. We now make this limit explicit and show that $(X_n)_{n \geq 0}$ converges toward $(1, \dots, 0)$ a.s. We start by noticing that $h_1(x) = x_1 \sum_{j \geq 2} x_j (p_1 - p_j) \geq 0$, which implies that :

$$X_n^1 \geq X_0^1 + \sum_{j=1}^n \gamma_j \rho_j \kappa_1 (X_{j-1}) + \sum_{j=1}^n \gamma_j \Delta M_{j,1}. \quad (2.5.41)$$

The martingale increment $\Delta M_{j,1}$ is bounded and a large enough C exists such that $\Delta M_{j,1} \leq \sqrt{C}$ a.s.. This implies that :

$$\left\| \sum_{j=1}^n \gamma_j \Delta M_{j,1} \right\|_{L^2}^2 \leq C \sum_{j=1}^n \gamma_j^2 \leq C \sup_{j \in \mathbb{N}} \left(\frac{\gamma_j}{\rho_j} \right) \sum_{j=1}^n \gamma_j \rho_j.$$

Since $\sum \rho_j \gamma_j = +\infty$, we can deduce that

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left(\frac{\sum_{j=1}^n \gamma_j \Delta M_{j,1}}{\sum_{j=1}^n \gamma_j \rho_j} \right)^2 = 0 \quad \text{so that} \quad \limsup_{n \rightarrow \infty} \frac{\sum_{j=1}^n \gamma_j \Delta M_{j,1}}{\sum_{j=1}^n \gamma_j \rho_j} \geq 0 \quad a.s.$$

Now, if we assume that $\omega \in \{X_{\infty,1} = 0\}$, we also deduce from the Toeplitz lemma that *a.s.*

$$\liminf_{n \rightarrow +\infty} \frac{1}{\sum_{j=1}^n \gamma_j \rho_j} \sum_{j=1}^n \gamma_j \rho_j \kappa_1(X_{j-1}(\omega)) \geq \inf_{(x_2, \dots, x_d) \in [0,1]^{d-1}, x_2 + \dots + x_d = 1} \kappa_1(0, x_2, \dots, x_d) > 0.$$

Thus, by (2.5.41), we deduce from what precedes that

$$\limsup_{n \rightarrow \infty} \frac{X_{n,1}(\omega)}{\sum_{j=1}^{n-1} \gamma_j \rho_j} > 0.$$

We obtain a contradiction with the boundedness of $(X_n)_{n \geq 1}$ and conclude that $\mathbb{P}(X_{\infty,1} = 0) = 0$. For (ii), we refer to [79] since the arguments here are similar. \square

2.5.2 Weak convergence of the normalized bandit (Proposition 2.3.2)

The proof of the weak convergence follows the lines of [79]. The idea is to prove the tightness of the pseudo-trajectories associated to the normalized sequence and to then show that any weak limit of this sequence is a solution of the martingale problem $(\mathcal{L}, \mathcal{C}_K^1(\mathbb{R}_+, (\mathbb{R}_+)^{d-1}))$ where \mathcal{L} is the infinitesimal generator defined in Proposition 2.3.2. Then, proving that uniqueness holds for the solutions of the martingale problem and for the invariant distribution, the convergence follows. Here, we choose to only detail the key step of the characterization of the limit. The rest of the proof can be obtained by a simple generalization of that of [79].

Proposition 2.5.1. Let $(p_1, \dots, p_d) \in (0, 1)^d$ with $p_d \leq \dots \leq p_2 < p_1$. Assume that $\gamma_n = \gamma_1 n^{-\frac{1}{2}}$ and $\rho_n = \rho_1 n^{-\frac{1}{2}}$. Set $Y_n = (Y_n^1, \dots, Y_n^d) = \rho_n^{-1} X_n$. Let f be a continuously differentiable function with compact support in \mathbb{R}_+^{d-1} and Lipschitz derivatives. We have

$$\mathbb{E}(f(Y_{n+1}^2, \dots, Y_{n+1}^d) - f(Y_n^2, \dots, Y_n^d) | \mathcal{F}_n) = \gamma_{n+1} (\mathcal{L}_d f(Y_n^2, \dots, Y_n^d) + o_{\mathbb{P}}(1)),$$

where \mathcal{L}_d is the PDMP generator defined in (2.3.15) and $\mathcal{F}_n = \sigma(Y_k, k \leq n)$.

Proof Set $\tilde{Y}_n = (Y_n^2, \dots, Y_n^d)$. An extension of Theorem 2.3.1 yields for any $(p_1, \dots, p_d) \in (0, 1)^d$ such that $p_1 > p_2 \geq \dots \geq p_d$,

$$\sup_{n \geq 1} \mathbb{E}[|\tilde{Y}_n|_1] < +\infty, \quad (2.5.42)$$

where $|u|_1 = \sum_{j=2}^d |u_j|$. Now, let us decompose the conditional expectation of the proposition as follows :

$$\mathbb{E}(f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n) | \mathcal{F}_n) = \sum_{i=1}^d \mathbb{E}((f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n)) 1_{\{I_{n+1}=i\}} | \mathcal{F}_n).$$

First, on the event $\{I_{n+1} = 1\}$, we have for every $j \geq 2$:

$$X_{n+1}^j = X_n^j(1 - \gamma_{n+1}1_{A_{n+1}^1}) + \gamma_{n+1}\rho_{n+1}\frac{X_n^1}{d-1}1_{C_{n+1}^\sigma},$$

where $C_{n+1}^\sigma = (A_{n+1}^1)^c \cup (A_{n+1}^1 \cap B_{n+1}^\sigma)$. Using that $\rho_n\rho_{n+1}^{-1} = O(\gamma_{n+1}^2)$, this yields for every $j \geq 2$

$$Y_{n+1}^j = Y_n^j(1 - \gamma_{n+1}1_{A_{n+1}^1}) + \gamma_{n+1}1_{C_{n+1}^\sigma} + O(\gamma_{n+1}^2|\tilde{Y}_n|_1).$$

Since f is Lipschitz continuous, we deduce that

$$\begin{aligned} \mathbb{E}\left((f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n))1_{\{I_{n+1}=i\}}|\mathcal{F}_n\right) &= X_n^1 p_1 (f(\tilde{Y}_n(1 - \gamma_{n+1})) - f(\tilde{Y}_n)) \\ &\quad + X_n^1 \frac{(1 - p_1 \sigma)}{d-1} (f(\tilde{Y}_n + \gamma_{n+1}v) - f(\tilde{Y}_n)) + O(\gamma_{n+1}^2|\tilde{Y}_n|_1), \end{aligned}$$

where $v = 1_{\mathbb{R}^{d-1}}$. Using again that f is Lipschitz continuous and that $1 - X_n^1 = \rho_n|\tilde{Y}_n|_1$, one obtains

$$\begin{aligned} \mathbb{E}\left((f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n))1_{\{I_{n+1}=i\}}|\mathcal{F}_n\right) &= p_1 (f(\tilde{Y}_n(1 - \gamma_{n+1})) - f(\tilde{Y}_n)) \\ &\quad + \frac{(1 - p_1 \sigma)}{d-1} (f(\tilde{Y}_n + \gamma_{n+1}v) - f(\tilde{Y}_n)) + O(\gamma_{n+1}^2(|\tilde{Y}_n|_1 + |\tilde{Y}_n|_1^2)). \end{aligned}$$

By the Taylor formula with order 1 and the fact that ∇f is Lipschitz continuous, the previous equality leads to

$$\mathbb{E}\left((f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n))1_{\{I_{n+1}=i\}}|\mathcal{F}_n\right) = \gamma_{n+1}(\nabla f(\tilde{Y}_n), -p_1\tilde{Y}_n + \frac{(1 - p_1 \sigma)}{d-1}v) + O(\gamma_{n+1}^2(1 + |\tilde{Y}_n|_1^2)),$$

where (\cdot, \cdot) stands for the usual scalar product on \mathbb{R}^{d-1} . By (2.5.42), one can readily check that $(\gamma_{n+1}(1 + |\tilde{Y}_n|_1^2))_{n \geq 1}$ converges to 0 in probability. As a consequence, one can conclude the case $I_{n+1} = 1$ by the following equality :

$$\mathbb{E}\left((f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n))1_{\{I_{n+1}=1\}}|\mathcal{F}_n\right) = \gamma_{n+1}(\mathcal{L}_d^1 f(\tilde{Y}_n) + o_{\mathbb{P}}(1)) \quad (2.5.43)$$

where for every $y \in \mathbb{R}^{d-1}$ and for every differentiable function $f : \mathbb{R}^{d-1} \mapsto \mathbb{R}$,

$$\mathcal{L}_d^1 f(y) = \sum_{j=2}^d \left(\frac{(1 - p_1 \sigma)}{d-1} - p_1 y_j \right) \partial_j f(y).$$

The above operator corresponds to the drift component of the limiting process. It remains now to exhibit the jump component. To this end, let us now consider the case $I_{n+1} = i$ with $i \in \{2, \dots, d\}$. We have

$$X_{n+1}^i = X_n^i + \gamma_{n+1}(1 - X_n^i)1_{A_{n+1}^i} + O(\gamma_{n+1}\rho_{n+1}X_n^i),$$

so that a normalization combined with similar arguments as previously yields

$$Y_{n+1}^i = Y_n^i + \frac{\gamma_{n+1}}{\rho_{n+1}}1_{A_{n+1}^i} + O(X_n^i) = Y_n^i + g1_{A_{n+1}^i} + O(X_n^i).$$

Now, for every $j \neq i$ ($j \geq 2$), we can remark that

$$Y_{n+1}^j = Y_n^j + O(X_n^j \vee X_n^i).$$

Using that f is Lipschitz continuous, we get for every $i \geq 2$,

$$\mathbb{E}\left((f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n))1_{\{I_{n+1}=i\}}|\mathcal{F}_n\right) = p_i X_n^i (f(Y_n + ge_i) - f(Y_n)) + O(\gamma_{n+1}^2 |\tilde{Y}_n|_1^2),$$

where $e_i = (0, \dots, 0, \underbrace{1}_i, 0, \dots, 0)'$. By (2.5.42), we can rewrite the previous equality as follows :

$$\forall i \geq 2, \quad \mathbb{E}\left((f(\tilde{Y}_{n+1}) - f(\tilde{Y}_n))1_{\{I_{n+1}=i\}}|\mathcal{F}_n\right) = \gamma_{n+1} \left(\frac{p_i}{g} Y_n^i (f(Y_n + ge_i) - f(Y_n)) + O_{\mathbb{P}}(1) \right)$$

and conclude the proof with a combination of this last statement with (2.5.43). □

2.6 Ergodicity of the PDMP

From now on, the variable $(X_t)_{t \geq 0}$ will refer to a trajectory of the PDMP associated with the normalized (over)-penalized bandit and bearing no relation to the multi-armed bandit sequence $(X_n)_{n \geq 1}$.

2.6.1 Wasserstein results

We begin the study of the ergodicity of the PDMP whose infinitesimal generator is (2.3.16) with some computations of the moments of the process.

Lemma 2.6.1. Let $(X_t)_{t \geq 0}$ be a Markov process, whose generator \mathcal{L} is defined by (2.3.16). If $\pi := b - cg > 0$, then $\sup \mathbb{E}[(X_t^x)^p] \leq C(1 + |x|^p)$. In particular, the invariant distribution π has moments of any order and

$$\forall t \geq 0 \quad \mathbb{E}(X_t) = \frac{a}{\pi} + \left(\mathbb{E}(X_0) - \frac{a}{\pi} \right) e^{-t\pi}$$

Proof Let us define $f_p(x) = x^p$. We have :

$$\begin{aligned} \mathcal{L}f_p(x) &= p(a - bx)x^{p-1} + cx((x + g)^p - x^p) \\ &= -p\pi f_p(x) + paf_{p-1}(x) + c \sum_{k=0}^{p-2} C_p^k g^{p-k} f_{k+1}(x), \end{aligned} \tag{2.6.44}$$

where we adopt the convention $\Sigma_\emptyset = 0$. If we now define $\alpha_p(t) = \mathbb{E}(X_t^p)$, the previous relation shows that α_p satisfies the ODE for any integer $p \geq 1$ defined by

$$\alpha_p(t)' + p\pi\alpha_p(t) = pa\alpha_{p-1}(t) + c \sum_{k=0}^{p-2} C_p^k g^{p-k} \alpha_{k+1}(t).$$

For example, with $p = 1$ we have $\alpha'_1(t) = -\pi\alpha_1(t) + a$, which implies that

$$\alpha_1(t) = \frac{a}{\pi} + \left(\mathbb{E}(X_0) - \frac{a}{\pi} \right) e^{-t\pi}.$$

The control of the moments of order $p > 1$ then follows from a recursion.

□

Rescaled two-armed bandit & Theorem 2.3.3

In the following, we will exploit Equation (2.6.44) to obtain a suitable upper bound of the Wasserstein distance \mathcal{W}_p between the law of X_t and the invariant measure μ_∞ of the PDMP. For this purpose, we note that the generator (2.3.16) possesses the stochastic monotonicity property, *i.e.*, a coupling (X, Y) exists starting from (x, y) (with $x > y$) such that $X_t \geq Y_t$ for any $t \geq 0$. The increase of the jump rate (with respect to the position) and the positivity of the jumps are of prime importance for this property. Such a coupling could be built as follows : we only allow simultaneous jumps of both components or a single jump of the highest one (see ([15]) for a similar procedure). The generator of this coupling (X, Y) starting from (x, y) with $x > y$ is given by :

$$\begin{aligned} f(x, y) = & (a - bx)\partial_x f(x, y) + (a - by)\partial_y f(x, y) \\ & + cy(f(x + g, y + g) - f(x, y)) + c(x - y)(f(x + g, y) - f(x, y)) \end{aligned} \quad (2.6.45)$$

with a symmetric expression when $y > x$. We now prove the main result.

Proof of theorem 2.3.3. Let μ_0 be a probability on \mathbb{R}_+^* and designate μ_∞ as the invariant distribution of the PDMP. Set

$$\mathcal{C}_t = \{ \nu \in \mathcal{P}(\mathbb{R}^2), \nu(dx \times \mathbb{R}_+) = \mu_t(dx), \nu(\mathbb{R}_+ \times dy) = \mu_\infty(dy) \}.$$

For any $\nu \in \mathcal{C}$, let $(X_t, Y_t)_{t \geq 0}$ denote the Markov process driven by (2.6.45) starting from ν . From the definition of \mathcal{W}_p and the stationarity of (Y_t) , we have for any t :

$$\mathcal{W}_p(\mu_t, \mu_\infty) \leq \inf \left\{ \nu \in \mathcal{C}_0, \left(\int_{\mathbb{R}_+^2} \mathbb{E}[|X_t^x - Y_t^y|^p] \nu(dx, dy) \right)^{\frac{1}{p}} \right\}.$$

At the price of a potential exchange of the coordinates, we can now work with some deterministic starting points x and y such that $x > y > 0$. Owing to the monotonicity of , we thus have for any $p \geq 1$

$$\mathbb{E}(|X_t^x - Y_t^y|^p) = \mathbb{E}(X_t^x - Y_t^y)^p.$$

Assume now that $p \in \mathbb{N}^*$, we observe that acts on $(x, y) \mapsto (x - y)^p$ as :

$$(x - y)^p = -p\pi(x - y)^p + pa(x - y)^{p-1} + c \sum_{k=0}^{p-2} C_p^k g^{p-k} (x - y)^{k+1}.$$

Setting $\beta_p(t) = \mathbb{E} |X_t^x - Y_t^y|^p$, we can immediately check that :

$$\dot{\beta}_p(t) + \pi p \beta_p(t) = \left(pa \beta_{p-1}(t) + c \sum_{k=0}^{p-2} C_p^k g^{p-k} \beta_{k+1}(t) \right). \quad (2.6.46)$$

When $p = 1$, (2.6.46) implies that : $\beta_1(t) = \beta_1(0)e^{-\pi t} \Rightarrow \mathbb{E}[X_t^x - Y_t^y] = (x - y)e^{-\pi t}$, so that :

$$\mathcal{W}_1(\mu_t, \mu_\infty) \leq \mathcal{W}_1(\mu_0, \mu_\infty) e^{-t\pi}.$$

For the lower-bound, we use :

$$\mathcal{W}_1(\mu_t, \mu_\infty) \geq \inf \left\{ \nu_t \in \mathcal{C}_t, \left| \int (x - y) \nu_t(dx, dy) \right| \right\} = |\mathbb{E}[X_t^{\mu_0}] - \mathbb{E}[Y_t^{\mu_\infty}]|,$$

which implies that :

$$\mathcal{W}_1(\mu_t, \mu_\infty) \geq \left| \int \mathbb{E}[X_t^x - Y_t^y] \mu_0(dx) \mu_\infty(dy) \right| = \left| \int (x - y) \mu_0(dx) \mu_\infty(dy) \right| e^{-\pi t}.$$

The lower-bound follows.

Now, let us consider the case $p > 1$ (with $p \in \mathbb{N}$). For $p = 2$, we have

$$(\beta_2(t)e^{2\pi t})' e^{-2\pi t} = (2a + cg^2)\beta_1(0)e^{-\pi t},$$

and an integration leads to $\beta_2(t)e^{2\pi t} - \beta_2(0) = \frac{2a+cg^2}{\pi}\beta_1(0)[e^{\pi t} - 1]$. As a consequence :

$$\beta_2(t) \leq e^{-2\pi t}\beta_2(0) + \frac{2a + cg^2}{\pi}\beta_1(0)e^{-\pi t}.$$

Using the inequalities $\sqrt{u+v} \leq \sqrt{u} + \sqrt{v}$ and $\beta_2 \geq \mathcal{W}_2^2$, we thus deduce that :

$$\mathcal{W}_2(\mu_t, \mu_\infty) \leq \mathcal{W}_2(\mu_0, \mu_\infty)e^{-\pi t} + \sqrt{\frac{2a + cg^2}{\pi}} \sqrt{\mathcal{W}_1(\mu_0, \mu_\infty)} e^{-\frac{\pi t}{2}}.$$

The result follows when $p = 2$ by setting :

$$\gamma_2 := \mathcal{W}_2(\mu_0, \mu_\infty) + \sqrt{\frac{2a + cg^2}{\pi}} \sqrt{\mathcal{W}_1(\mu_0, \mu_\infty)}.$$

A recursive argument based on (2.6.46) shows that a constant γ_p exists that only depends on μ_0 and μ_∞ such that :

$$\mathcal{W}_p(\mu_t, \mu_\infty) \leq \gamma_p e^{-\frac{\pi}{p}t}.$$

□

2.6.2 Proof of total variation results

As mentioned before, the idea is to wait until the paths get close (with a probability controlled by the Wasserstein bound) and then to try to stick them (with high probability). Since the jump

size is deterministic, sticking the paths implies a non trivial coupling of the jump times which is described in the lemma below.

We begin by establishing the next useful lemma.

Lemma 2.6.2. Let $\varepsilon > 0$ and $t \geq \frac{1}{b} \ln(1 + \varepsilon)$. A coupling $(X_t, Y_t)_{t \geq 0}$ of paths driven by (2.3.16) exists such that on $A_{x_0, \varepsilon}$:

$$\mathbb{P}(X_t = Y_t, t \geq s) \geq \left(1 - \frac{c}{b} x_0 \varepsilon - e^{-\frac{a}{b} cs} - \frac{c\varepsilon}{b}\right) \max(0, 1 - \frac{c}{b} \varepsilon(x_0 + g)),$$

where $A_{x_0, \varepsilon} = \{(x, y) | \frac{a}{b} < x \leq x_0, 0 < x - y \leq \varepsilon\}$.

Proof Let $\varepsilon > 0$ and $(x, y) \in A_{x_0, \varepsilon}$ (in particular, $x > y$). Denote by (X_t^x) and (X_t^y) the paths starting from x and y respectively, T_1^x and T_1^y the associated first jump times, and by T_2^x the second jump time of (X_t^x) . It can be noticed that :

$$\mathbb{P}(X_t = Y_t, t \geq s) \geq \mathbb{P}(X_{T_1^y}^x = X_{T_1^y}^y, T_1^y \leq s).$$

We aim to build a coupling that leads to a sharp lower-bound of the r.h.s. For this purpose, note that if $T_1^x < T_1^y < T_2^x$, the triple (T_1^x, T_1^y, T_2^x) satisfies :

$$X_{T_1^y}^y = X_{T_1^y}^x \iff \frac{a}{b} + \left(y - \frac{a}{b}\right) e^{-bT_1^y} + g = \frac{a}{b} + \left(X_{T_1^x}^x - \frac{a}{b}\right) e^{-b(T_1^y - T_1^x)}.$$

Considering that $X_{T_1^x}^x = \frac{a}{b} + (x - \frac{a}{b})e^{-T_1^x} + g$ and defining $\psi(t) = \frac{1}{b} \ln \left(e^{bt} + \frac{x-y}{g}\right)$, we can verify that $X_{T_1^y}^y = X_{T_1^y}^x \leq s$ and $T_1^x < T_1^y < T_2^x$ as soon as

$$T_1^y = \psi(T_1^x) \leq s \quad \text{and} \quad T_2^x \geq \psi(T_1^x),$$

since $\psi(t) \geq t$. We are naturally encouraged to consider $S_1^{x,s} = \psi(T_1^x) 1_{\{\psi(T_1^x) \leq s\}}$ and it is well known that the law of (T_1^x, T_1^y) can be described through the maximal coupling :

$$T_1^y = \Theta U + (1 - \Theta)V_y, \quad \psi(T_1^x) = \Theta U + (1 - \Theta)V_x,$$

where V_x, V_y, Θ and U are independent, $U \sim \frac{\mathbb{P}_{T_1^y} \wedge \mathbb{P}_{\psi(T_1^x)}}{\|\mathbb{P}_{S_1^{x,s}} \wedge \mathbb{P}_{T_1^y}\|_{TV}}$ and $\Theta \sim \mathcal{B}(p)$ where $p = \|\mathbb{P}_{S_1^{x,s}} \wedge \mathbb{P}_{T_1^y}\|_{TV}$. With this coupling, if $q(t, z) = \mathbb{P}(T_1^z \geq \psi(t) - t)$, the Strong Markov property yields

$$\mathbb{P}(T_2^x - T_1^x | (T_1^x, T_1^y)) = \mathbb{P}(T_2^x \geq \psi(T_1^x) | T_1^x) = q(T_1^x, X_{T_1^x}^x).$$

Since $z \mapsto q(t, z)$ is increasing and $x > a/b$ (from the assumption on $A_{x_0, \varepsilon}$), we deduce that $X_{T_1^x}^x \leq x + g$ and it therefore follows that :

$$\mathbb{P}(T_2^x > T_1^y | (T_1^x, T_1^y)) \geq q(t, x + g) \geq q(0, x + g).$$

given that $t \mapsto \psi(t) - t$ is a non-decreasing function. As a consequence, we obtain that with this coupling :

$$\mathbb{P}(X_{T_1^y}^x = X_{T_1^y}^y, T_1^y \leq s) \geq q(0, x + g) \mathbb{P}(\Theta = 1) = q(0, x + g) \|\mathbb{P}_{S_1^{x,s}} \wedge \mathbb{P}_{T_1^y}\|_{TV}. \quad (2.6.47)$$

It remains to find a lower bound of the total variation distance involved in the r.h.s. of the above inequality . Recall that

$$\|\mathbb{P}_{S_1^{x,s}} \wedge \mathbb{P}_{T_1^y}\|_{TV} = \int_0^{+\infty} f_y(t) \wedge g_{x,s}(t) dt,$$

where f_y and $g_{x,s}$ denote the densities of T_1^y and $S_1^{x,s}$, respectively. We therefore have :

$$\forall t > 0, \quad f_y(t) = c\phi(y,t)e^{-\int_0^t c\phi(y,u)du} \quad \text{with} \quad \phi(y,t) = \frac{a}{b} + (y - \frac{a}{b})e^{-bt},$$

and a change of variable yields :

$$\forall t > 0, \quad g_x(t) = f_x(\psi^{-1}(t))(\psi^{-1})'(t)1_{\{\psi(0) \leq t \leq s\}}. \quad (2.6.48)$$

On the one hand, since $(x,y) \in A_{x_0,\epsilon}$, we can check that :

$$\forall t \geq 0, \quad \phi(x,t) - \epsilon e^{-bt} \leq \phi(y,t) \leq \phi(x,t),$$

and we can then conclude that :

$$\forall t > 0, \quad f_y(t) \geq f_x(t) - \epsilon e^{-bt}.$$

On the other hand, note that :

$$\forall t > \psi(0), \quad \psi^{-1}(t) = \frac{1}{b} \ln \left(e^{bt} - \frac{x-y}{g} \right) \leq t \quad \text{and} \quad (\psi^{-1})'(t) = \frac{e^{bt}}{e^{bt} - \frac{x-y}{g}} \geq 1,$$

and we can deduce from (2.6.48) that $\forall t \in [\psi(0), s]$:

$$g_x(t) \geq c\phi(x, \psi^{-1}(t))e^{-\int_0^t c\phi(x,s)ds} \geq c\phi(x,t)e^{-\int_0^t c\phi(x,s)ds} = f_x(t).$$

Note that we used that $t \mapsto \phi(x,t)$ is decreasing since $x > a/b$. Thus,

$$\left(\mathbb{P}_{T_1^y} \wedge \mathbb{P}_{S_1^x} \right) (dt) \geq h(t)dt \quad \text{with} \quad h(t) = (f_x(t) - \epsilon e^{-bt})1_{\psi(0) \leq t \leq s} dt.$$

As a consequence,

$$\|\mathbb{P}_{S_1^{x,s}} \wedge \mathbb{P}_{T_1^y}\|_{TV} \geq e^{-\int_0^{\psi(0)} c\phi(x,u)du} - e^{-\int_0^s c\phi(x,u)du} - \frac{\epsilon}{b}.$$

Checking that $\psi(0) \leq \epsilon/b$ and that $\forall t \geq 0, a/b \leq \phi(x,t) \leq x \leq x_0$, we deduce that

$$\|\mathbb{P}_{S_1^{x,s}} \wedge \mathbb{P}_{T_1^y}\|_{TV} \geq e^{-\frac{cx_0\epsilon}{b}} - e^{-\frac{a}{b}cs} - \frac{\epsilon}{b} \geq 1 - \frac{cx_0\epsilon}{b} - e^{-\frac{a}{b}cs} - \frac{\epsilon}{b},$$

where we used $e^{-u} \geq 1 - u$ for $u \geq 0$ in the second line. To conclude the proof, it remains to plug this inequality into (2.6.47) and to observe that :

$$q(0, x+g) \geq q(0, x_0+g) = e^{-\int_0^{\psi(0)} c\phi(x_0+g,s)ds} \geq 1 - c\psi(0)(x_0+g) \geq 1 - \frac{c}{b}\epsilon(x_0+g).$$

□

We now provide the proof of the ergodicity w.r.t. the total variation distance.

Proof of Theorem 2.3.4. For any starting distribution μ_0 ,

$$\|\mu_0 P_t - \mu_\infty\|_{TV} \leq \int \|\delta_x P_t - \delta_y P_t\|_{TV} \mu_0(dy) \mu_\infty(dx). \quad (2.6.49)$$

The idea is to use the Wasserstein coupling during a time t_1 and to then try to stick the paths on the interval $[t_1, t]$ using Lemma 2.6.2. Consider $A_{x_0, \varepsilon}$ defined in Lemma 2.6.2 and the alternative set $A_{x_0, \varepsilon}^* = \{(x, y), a/b < y < x_0, 0 < y - x \leq \varepsilon\}$. Set $B_{x_0, \varepsilon} = A_{x_0, \varepsilon} \cup A_{x_0, \varepsilon}^*$, we have :

$$1 - \|\delta_x P_t - \delta_y P_t\|_{TV} \geq \mathbb{P}(X_t^x = Y_t^y | (X_{t_1}^x, Y_{t_1}^y) \in B_{x_0, \varepsilon}) \mathbb{P}((X_{t_1}^x, Y_{t_1}^y) \in B_{x_0, \varepsilon}). \quad (2.6.50)$$

Since the Wasserstein coupling preserves the order and since $x > a/b$ $\mu_\infty(dx)$ -a.s., it can be noted that $\mu_\infty(dx)$ -a.s.,

$$(X_{t_1}^x, Y_{t_1}^y) \in B_{x_0, \varepsilon} \iff \begin{cases} X_{t_1}^x - X_{t_1}^y \leq \varepsilon \text{ and } X_{t_1}^x \leq x_0 & \text{if } x \geq y \\ X_{t_1}^y - X_{t_1}^x \leq \varepsilon \text{ and } X_{t_1}^y \leq x_0 & \text{if } x < y. \end{cases}$$

It follows that for every $p > 0$, $\mu_\infty(dx)$ almost surely :

$$\begin{aligned} \mathbb{P}((X_{t_1}^x, Y_{t_1}^y) \in B_{x_0, \varepsilon}^c) &\leq \mathbb{P}(|X_{t_1}^x - X_{t_1}^y| > \varepsilon) + \mathbb{P}(X_{t_1}^x > x_0) + \mathbb{P}(X_{t_1}^y > x_0) \\ &\leq \frac{1}{\varepsilon} \mathbb{E}[|X_{t_1}^x - X_{t_1}^y|] + \frac{1}{x_0^p} (\mathbb{E}[(X_{t_1}^x)^p] + \mathbb{E}[(X_{t_1}^y)^p]). \end{aligned}$$

On the basis of Theorem 2.3.3 and Lemma 2.6.1, a constant C_p exists such that C_p depends on p , μ_0 and μ_∞ but not on t_1 and satisfies :

$$\int \mathbb{P}((X_{t_1}^x, Y_{t_1}^y) \in B_{x_0, \varepsilon}^c) \mu_0(dy) \mu_\infty(dx) \leq \frac{\mathcal{W}_1(\mu_0, \mu_\infty)}{\varepsilon} e^{-\pi t_1} + \frac{C_p}{x_0^p}.$$

Finally, Lemma 2.6.2 leads to :

$$\begin{aligned} &\mathbb{P}(X_t^x = Y_t^y | (X_{t_1}^x, Y_{t_1}^y) \in B_{x_0, \varepsilon}) \\ &\geq \left(1 - \frac{c}{b} x_0 \varepsilon - e^{-\frac{a}{b} c(t-t_1)} - \frac{c\varepsilon}{b}\right) \left\{0 \vee 1 - \frac{c}{b} \varepsilon (x_0 + g)\right\} \end{aligned}$$

so that by plugging the previous inequalities into (2.6.50) and (2.6.49), it can be deduced that for every $p > 1$, a constant \tilde{C}_p exists such that for every $t \geq 0$, for every x_0 and ε such that $x_0 \varepsilon \leq b/2c$ (with $x_0 > 1$ and $\varepsilon \in (0, 1)$),

$$\|\mu_0 P_t - \mu_\infty\|_{TV} \leq \tilde{C}_p \left(x_0 \varepsilon + e^{-\frac{a}{b} c(t-t_1)} + \varepsilon + \frac{1}{\varepsilon} e^{-\pi t_1} + \frac{1}{x_0^p} \right).$$

If we try to optimize the above bound, we set $t_1 = \delta t$, $x_0 = C_1 e^{\alpha t}$, $\varepsilon = C_2 e^{-\beta t}$ with $\delta \in (0, 1)$ and $\beta > \alpha > 0$ and deduce that a constant \check{C}_p exists such that :

$$\|\mu_0 P_t - \mu_\infty\|_{TV} \leq \check{C}_p \exp \left(-t \left\{ \beta - \alpha \wedge \frac{ca}{b} (1 - \delta) \wedge \delta \pi - \beta \wedge \alpha p \right\} \right).$$

We can choose p as large as we want (μ_0 has moments of any order) and thus α arbitrarily small. The result then follows using an optimization on (β, δ) . \square

2.7 Technical result for the pseudo-regret upper bound

Lemma 2.7.1. Let $\alpha > 0$, $\gamma_1 \in (0, 1)$ and $\tilde{n} \in \mathbb{N}$ such that $\alpha\gamma_{\tilde{n}} < 1$ and $\tilde{n} \geq 1/(\alpha\gamma_1)^2$. We have :

$$\forall n \geq \tilde{n} \quad \sum_{j=\tilde{n}}^{n-1} \gamma_j \prod_{l=j}^{n-1} (1 - \alpha\gamma_l) \leq \frac{1}{\alpha}$$

Proof Let $j \geq \tilde{n}$. On the basis of the inequality $\ln(1+x) \geq x$ for $x > -1$, we have

$$\prod_{l=j}^{n-1} (1 - \alpha\gamma_l) = \exp \left(\ln \sum_{l=j}^{n-1} (1 - \alpha\gamma_l) \right) \leq \exp \left(- \sum_{l=j}^{n-1} \alpha\gamma_l \right)$$

Using that $x \mapsto 1/\sqrt{x}$ is decreasing,

$$\sum_{l=j}^{n-1} \gamma_l = \sum_{l=j}^{n-1} \frac{\gamma_1}{\sqrt{l}} \geq \gamma_1 \sum_{l=j}^{n-1} \int_l^{l+1} \frac{1}{\sqrt{x}} dx = \gamma_1 \int_j^n \frac{1}{\sqrt{x}} dx = 2\gamma_1(\sqrt{n} - \sqrt{j})$$

so that :

$$\sum_{j=n_0}^{n-1} \gamma_j \prod_{l=j}^{n-1} (1 - \alpha\gamma_l) \leq \gamma_1 e^{-2\alpha\gamma_1\sqrt{n}} \sum_{j=n_0}^{n-1} \frac{e^{2\alpha\gamma_1\sqrt{j}}}{\sqrt{j}}.$$

Checking that $x \mapsto \frac{1}{\sqrt{x}} e^{\alpha\gamma_1\sqrt{x}}$ is non-decreasing on $[\frac{1}{(\alpha\gamma_1)^2}, \infty)$ it can be deduced that for any $j \geq n_0$,

$$\sum_{j=n_0}^{n-1} \frac{1}{\sqrt{j}} e^{2\alpha\gamma_1\sqrt{j}} \leq \int_{n_0}^n \frac{1}{\sqrt{x}} e^{2\alpha\gamma_1\sqrt{x}} dx \leq \frac{1}{\alpha\gamma_1}.$$

The lemma follows.

Chapitre 3

Stochastic Heavy Ball

Sommaire

3.1	Introduction	112
3.2	Stochastic Heavy Ball	113
3.2.1	Deterministic Heavy Ball	113
3.2.2	Stochastic HBF	114
3.2.3	Baseline assumptions	115
3.2.4	Main results	118
3.3	Almost sure convergence of the stochastic heavy ball	121
3.3.1	Preliminary result	121
3.3.2	Convergence to a local minimum	126
3.4	Convergence rates for strongly convex functions	136
3.4.1	Quadratic case	136
3.4.2	The non-quadratic case under exponential memory	142
3.5	Limit of the rescaled algorithm	145
3.5.1	Rescaling stochastic HBF	145
3.5.2	Tightness	146
3.5.3	Identification of the limit	148
3.5.4	Limit variance	152
3.6	Numerical experiments	155
3.7	Technical results	158
3.7.1	Standard tools of stochastic algorithms	158
3.7.2	Expectation of the supremum of the square of sub-Gaussian random variables	162

Abstract This paper deals with a natural stochastic optimization procedure derived from the so-called Heavy-ball method differential equation, which was introduced by Polyak in the 1960s with his seminal contribution [109]. The Heavy-ball method is a second-order dynamics that was investigated to minimize convex functions f . The family of second-order methods recently received a large amount of attention, until the famous contribution of Nesterov [96], leading to the explosion of large-scale optimization problems. This work provides an in-depth description of the stochastic heavy-ball method, which is an adaptation of the deterministic one when only unbiased evaluations of the gradient are available and used throughout the iterations of the algorithm. We first describe some almost sure convergence results in the case of general non-convex coercive functions f . We then examine the situation of convex and strongly convex potentials and derive some non-asymptotic results about the stochastic heavy-ball method. We end our study with limit theorems on several rescaled algorithms.

3.1 Introduction

Finding the minimum of a function f over a set Ω with an iterative procedure is very popular among numerous scientific communities and has many applications in optimization, image processing, economics and statistics, to name a few. We refer to [95] for a general survey on optimization algorithms and discussions related to complexity theory, and to [97, 32] for a more focused presentation on convex optimization problems and solutions. The most widespread approaches rely on some first-order strategies, with a sequence $(X_k)_{k \geq 0}$ that evolves over Ω with a first-order recursive formula $X_{k+1} = \Psi[X_k, f(X_k), \nabla f(X_k)]$ that uses a local approximation of f at point X_k , where this approximation is built with the knowledge of $f(X_k)$ and $\nabla f(X_k)$ alone. Among them, we refer to the steepest descent strategy in the convex unconstrained case, and to the Frank-Wolfe [50] algorithm in the compact convex constrained case. A lot is known about first-order methods concerning their rates of convergence and their complexity. In comparison to second-order methods, first-order methods are generally slower and are significantly degraded on ill-conditioned optimization problems. However, the complexity of each update involved in first-order methods is relatively limited and therefore useful when dealing with a large-scale optimization problem, which is generally expensive in the case of Interior Point and Newton-like methods. A second-order “optimal” method was proposed in [96] in the 1980s’ (also see [16] for an extension of this method with proximal operators). The so-called *Nesterov Accelerated Gradient Descent* (NAGD) has particularly raised considerable interest due to its numerical simplicity, to its low complexity and to its mysterious behavior, making this method very attractive for large-scale machine learning problems. Among the available interpretations of NAGD, some recent advances have been proposed concerning the second-order dynamical system by [114], being a particular case of the generalized Heavy Ball with Friction method (referred to as HBF in the text), as previously pointed out in [35, 36]. In particular, as highlighted in [35], NAGD may be seen as a specific case of HBF after a time rescaling $t = \sqrt{s}$, thus making the acceleration explicit through this change of variable, as well as being closely linked to the modified Bessel functions when f is quadratic.

A growing field of interest related to these optimization algorithms concerns the development of efficient procedures when only noisy gradients are available at each iteration of the procedure. On the practical side, this question was first introduced in the seminal contributions on stochastic approximation and optimization of [112] and [69]. Even though the Robbins-Monro algorithm is able to achieve an optimal $O(1/n)$ rate of convergence for strongly convex functions, its ability is highly sensitive to the step sizes used. This remark led [110] to develop an averaging method that

makes it possible to use longer step sizes of the Robbins-Monro algorithm, and to then average these iterates with a Cesaro procedure so that this method produces optimal results in the minimax sense (see [95]) for convex and strongly convex minimization problems, as pointed out in [12].

On the theoretical side, numerous studies have addressed a dynamical system point of view and studied the close links between stochastic algorithms and their deterministic counterparts for some general function f (*i.e.*, even non convex). These links originate in the famous Kushner-Clark Theorem (see [73]) and successful improvements have been obtained using differential geometry by [22, 20] on the long-time behavior of stochastic algorithms. In particular, a growing field of interest concerns the behavior of self-interacting stochastic algorithms (see, among others, [24] and [53]) because these non-Markovian processes produce interesting features from the modeling point of view (an illustration may be found in [53]).

Several theoretical contributions to the study of second-order stochastic optimization algorithms exist. [80] explores some adaptations of the NAGD in the stochastic case for composite (strongly or not) convex functions. Other authors [57, 58] obtained convergence results for the stochastic version of a variant of NAGD for non-convex optimization for gradient Lipschitz functions but these methods cannot be used for the analysis of the Heavy-ball algorithm. Finally, a recent work [122] proposes a unified study of some stochastic momentum algorithms while assuming restrictive conditions on the noise of each gradient evaluation and on the constant step size used. It should be noted that [122] provides a preliminary result on the behavior of the stochastic momentum algorithms in the non-convex case with possible multi-well situations. Our work aims to study the properties of a stochastic optimization algorithm naturally derived from the generalized heavy ball with friction method.

Our paper is organized as follows : Section 3.2 introduces the stochastic algorithm as well as the main assumptions needed to obtain some results on this optimization algorithm. For the sake of readability, these results are then provided in Section 3.2.4 without too many technicalities. The rest of the paper then deals with the proof of these results. Section 3.3 is dedicated to the almost sure convergence result we can obtain in the case of a non-convex function f with several local minima. Section 3.4 establishes the convergence rates of the stochastic heavy ball in the strongly convex case. Section 3.5 provides a central limit theorem in a particular case of the algorithm. Appendix 3.7 consists of some important results on the supremum of certain random variables needed for the non-convex case.

3.2 Stochastic Heavy Ball

We begin with a brief description of what is known about the underlying dynamical system.

3.2.1 Deterministic Heavy Ball

This method introduced by Polyak in [109] is inspired from the physical idea of producing some inertia on the trajectory to speed up the evolution of the underlying dynamical system : a ball evolves over the graph of a function f and is submitted to both damping (due to a friction on the graph of f) and acceleration. More precisely, this method is a second-order dynamical system described by the following O.D.E. :

$$\ddot{x}_t + \gamma_t \dot{x}_t + \nabla f(x_t) = 0, \quad (3.2.1)$$

where $(\gamma_t)_{t \geq 0}$ corresponds to the damping coefficient, which is a key parameter of the method. In particular, it is shown in [35] that the trajectory converges only under some restrictive conditions on the function $(\gamma_t)_{t \geq 0}$, namely :

- if $\int_0^{+\infty} \gamma_s ds = \infty$, then $(f(x_t))_{t \geq 0}$ converges,
- if $\int_0^{\infty} e^{-\int_0^t \gamma_s ds} dt < \infty$, then $(x_t)_{t \geq 0}$ converges towards one of the minima of any convex function f .

Intuitively, these conditions translate the oscillating nature of the solutions of (3.2.1) into a quantitative setting for the convergence of the trajectories : if $\gamma_t \rightarrow 0$ is sufficiently fast, then the trajectory cannot converge (the limiting case being $\ddot{x} + \nabla f(x) = 0$). These properties lead us to consider two natural families of functions $(\gamma_t)_{t \geq 0}$: $\gamma_t = r/t$ with $r > 1$ and $\gamma_t = \gamma > 0$. To convert (3.2.1) into a tractable iterative algorithm, it is necessary to rewrite this O.D.E. using a coupled momentum equation. Consistent with [36], (3.2.1) is *equivalent* to the following integro-differential equation :

$$\dot{x}_t = -\frac{1}{k(t)} \int_0^t h(s) \nabla f(x_s) ds, \quad (3.2.2)$$

where h and k are two memory functions related to γ . In the natural situation of two positive increasing functions h and k , if $(x_t)_{t \geq 0}$ is a solution of (3.2.2), then $(\tilde{x}_s)_{s \geq 0}$ is solution of (3.2.1) with :

$$\tilde{x}_s = x_{\tau(s)} \quad \text{and} \quad \dot{\tau}(s) = \sqrt{(kh^{-1})(\tau(s))} \quad \text{with} \quad \gamma_s = \frac{\dot{k}h + k\dot{h}}{2h^{3/2}k^{1/2}} \circ \tau(s).$$

We can consider two typical situations where the deterministic HBF (3.2.1) converges (see [35] for further details) :

- The *exponentially memoried HBF* corresponds to the choice $k(t) = \lambda e^{\lambda t}$ and $h(t) = e^{\lambda t}$ and to a constant damping function $\gamma_s = \sqrt{\lambda}$ when the time scale is given by $\tau(s) = \sqrt{\lambda} s$. Note that in this situation, the two convergence conditions are satisfied since :

$$\int_0^{+\infty} \gamma_s ds = \int_0^{+\infty} \sqrt{\lambda} ds = +\infty \quad \text{and} \quad \int_0^{\infty} e^{-\int_0^t \gamma_s ds} dt = \int_0^{\infty} e^{-\sqrt{\lambda} t} dt < \infty.$$

- The *polynomially memoried HBF* corresponds to the choice $k(t) = t^{\alpha+1}$ and $h(t) = (\alpha+1)t^\alpha$ and is associated with an asymptotically vanishing damping $\gamma_s = \frac{2\alpha+1}{s}$ and a time scale $\tau(s) = \frac{s^2}{4(\alpha+1)}$, where the choice $\alpha = 1$ is associated with the NAGD (see [114] and their “magic” constant $3 = 2\alpha + 1$ in that case).

3.2.2 Stochastic HBF

All these remarks lead to the consideration of a natural stochastic version of (3.2.2) when $h = \dot{k}$. As pointed out by [53], the introduction of an auxiliary function $y_t = k(t)^{-1} \int_0^t h(s) \nabla f(x_s) ds$ makes it possible to obtain a first-order Markov evolution because $\dot{y}_t = r_t(\nabla f(x_t) - y_t)$ with $r_t = \frac{\dot{h}(t)}{k(t)}$. Hence, we define the stochastic Heavy Ball system as $(X_0, Y_0) = (x, y) \in \mathbb{R}^{2d}$ and :

$$\begin{cases} X_{n+1} = X_n - \gamma_{n+1} Y_n \\ Y_{n+1} = Y_n + \gamma_{n+1} r_n (\nabla f(X_n) - Y_n) + \gamma_{n+1} r_n \Delta M_{n+1}, \end{cases} \quad (3.2.3)$$

where the natural filtration of the sequence $(X_n, Y_n)_{n \geq 0}$ is denoted $(\mathcal{F}_n)_{n \geq 1}$ and :

- (ΔM_n) is a sequence of \mathcal{F}_n -martingale increments. For applications, ΔM_{n+1} usually represents the difference between the “true” value of $\nabla f(X_n)$ and the one observed at iteration n denoted $\partial_x F(X_n, \xi_n)$, where $(\xi_n)_n$ is a sequence of i.i.d. random variables and F is an \mathbb{R}^d -valued measurable function such that :

$$\forall u \in \mathbb{R}^d \quad \mathbb{E} [\partial_x F(u, \xi)] = \nabla f(u)$$

In this case,

$$\Delta M_{n+1} = \nabla f(X_n) - \partial_x F(X_n, \xi_n). \quad (3.2.4)$$

The randomness appears in the second component of the algorithm (3.2.3), whereas it was handled in the first component in [53]. We will introduce some assumptions on f and on the martingale sequence later.

- $(\gamma_n)_{n \geq 1}$ corresponds to the step size used in the stochastic algorithm, associated with the “time” of the algorithm represented by :

$$\Gamma_n = \sum_{k=1}^n \gamma_k \quad \text{such that} \quad \lim_{n \rightarrow +\infty} \Gamma_n = +\infty.$$

For the sake of convenience, we also define :

$$\Gamma_n^{(2)} = \sum_{k=1}^n \gamma_k^2,$$

which may converge or not according to the choice of the sequence $(\gamma_k)_{k \geq 1}$.

- $(r_n)_{n \geq 1}$ is a deterministic sequence that mimics the function $t \mapsto r_t$ defined as :

$$r_n = \frac{h(\Gamma_n)}{k(\Gamma_n)}. \quad (3.2.5)$$

In particular, when an exponentially weighted HBF with $k(t) = e^{rt}$ is chosen, we have $r_n = r > 0$, regardless of the value of n . In the other situation where $k(t) = t^r$, we obtain $r_n = r \Gamma_n^{-1}$.

3.2.3 Baseline assumptions

We introduce some of the general assumptions we will work with below. Some of these conditions are very general, whereas others are more specifically dedicated to the analysis of the strongly convex situation. We will use the notation $\|\cdot\|$ (resp. $\|\cdot\|_F$) below to refer to the Euclidean norm on \mathbb{R}^d (resp. the Frobenius norm on $\mathcal{M}_{d,d}(\mathbb{R})$). Finally, when $A \in \mathcal{M}_{d,d}(\mathbb{R})$, $\|A\|_\infty$ will refer to the maximal size of the modulus of the coefficients of A : $\|A\|_\infty := \sup_{i,j} |A_{i,j}|$. Our theoretical results will obviously not involve all of these hypotheses simultaneously.

Function f We begin with a brief enumeration of assumptions on the function f .

- Assumption (\mathbf{H}_s) : f is a function in $\mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ such that :

$$\lim_{|x| \rightarrow +\infty} f(x) = +\infty \quad \text{and} \quad \|D^2 f\|_\infty := \sup_{x \in \mathbb{R}^d} \|D^2 f(x)\|_F < +\infty \quad \text{and} \quad \|\nabla f\|^2 \leq c_f f.$$

The assumption (\mathbf{H}_s) is weak : it essentially requires that f be smooth, coercive and have, at the most, a quadratic growth on ∞ . In particular, no convexity hypothesis is made when f satisfies (\mathbf{H}_s) . It would be possible to extend most of our results to the situation where f is L -smooth (with a L -Lipschitz gradient), but we preferred to work with a slightly more stringent condition to avoid additional technicalities.

- Assumption $(\mathbf{H}_{SC}(\alpha))$: f is a convex function such that $\alpha = \inf_{x \in \mathbb{R}^d} \text{Sp}(D^2 f(x)) > 0$ and $D^2 f$ is Lipschitz.

In particular, $(\mathbf{H}_{SC}(\alpha))$ implies that f is α -strongly convex, meaning that :

$$\forall (x, y) \in \mathbb{R}^d \times \mathbb{R}^d \quad f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle + \frac{\alpha}{2} \|x - y\|^2.$$

Of course, $(\mathbf{H}_{SC}(\alpha))$ is still standard and is the most favorable case when dealing with convex optimization problems, leading to the best possible achievable rates. $(\mathbf{H}_{SC}(\alpha))$ translates the fact that the spectrum of the Hessian matrix at point x , denoted by $\text{Sp}(D^2 f(x))$, is lower bounded by $\alpha > 0$, uniformly over \mathbb{R}^d . The fact that $D^2 f$ is assumed to be Lipschitz will be useful to achieve convergence rates in Section 3.4.2.

Noise sequence $(\Delta M_{n+1})_{n \geq 1}$ We will essentially use three types of assumptions alternatively on the noise of the stochastic algorithm (3.2.3). The first and second assumptions are concerned with a concentration-like hypothesis. The first one is very weak and asserts that the noise has a bounded \mathbb{L}^2 norm.

- Assumption $(\mathbf{H}_{\sigma,p})$: ($p \geq 1$) For any integer n , we have :

$$\mathbb{E}(\|\Delta M_{n+1}\|^p | \mathcal{F}_n) \leq \sigma^2 (1 + f(X_n))^p.$$

The assumption $(\mathbf{H}_{\sigma,2})$ is a standard convergence assumption for general stochastic algorithms. For some non-asymptotic rates of convergence results, we will rely on $(\mathbf{H}_{\sigma,p})$ for any $p \geq 1$. In this case, we will denote the assumption by $(\mathbf{H}_{\sigma,\infty})$. Finally, let us note that the condition could be slightly alleviated by replacing the right-hand member by $\sigma^2 (1 + f(X_n) + |Y_n|^2)^p$. However, in view of the standard case (3.2.4), this improvement has little interest in practice, which explains our choice.

- Assumption $(\mathbf{H}_{Gauss,\sigma})$: For any integer n , the Laplace transform of the noise satisfies :

$$\forall t \geq 0 \quad \mathbb{E}[\exp(t \Delta M_{n+1}) | \mathcal{F}_n] \leq e^{\frac{\sigma^2 t^2}{2}}.$$

This hypothesis is much stronger than $(\mathbf{H}_{\sigma,p})$ and translates a sub-Gaussian behavior of $(\Delta M_{n+1})_{n \geq 1}$. In particular, it can be easily shown that $(\mathbf{H}_{Gauss,\sigma})$ implies $(\mathbf{H}_{\sigma,p})$. Hence, $(\mathbf{H}_{Gauss,\sigma})$ is somewhat restrictive and will be used only to obtain one important result in the non-convex situation for the almost sure limit of the stochastic heavy ball with multiple wells.

- Assumption $(\mathbf{H}_{\mathcal{E}})$: For any iteration n , the noise of the stochastic algorithm satisfies :

$$\forall v \in \mathcal{S}_{\mathbb{R}^d}^1 \quad \mathbb{E}(|\langle \Delta M_n, v \rangle| \mid X_n, Y_n) \geq c_v > 0.$$

This assumption will be essential to derive an almost sure convergence result towards minimizers of f . Roughly speaking, this assumption states that the noise is uniformly elliptic given any current position of the algorithm at step n : the projection of the noise has a non-vanishing component over all directions v . We will use this assumption to guarantee the ability of (3.2.3) to get out of any unstable point.

Step sizes One important step in the use of stochastic minimization algorithms relies on an efficient choice of the step sizes involved in the recursive formula (*e.g.* in Equation 3.2.3). We will deal with the following sequences $(\gamma_n)_{n \geq 0}$ below.

- Assumption $(\mathbf{H}_{\beta}^{\gamma})$: The sequence $(\gamma_n)_{n \geq 0}$ satisfies :

$$\forall n \in \mathbb{N} \quad \gamma_n = \frac{\gamma}{n^{\beta}} \quad \text{with} \quad \beta \in (0, 1],$$

leading to :

$$\forall \beta \in (0, 1) \quad \Gamma_n \sim \frac{\gamma}{1-\beta} n^{1-\beta} \quad \text{whereas} \quad \Gamma_n \sim \gamma \log n \quad \text{when} \quad \beta = 1.$$

Memory size We consider the exponentially and polynomially-weighted HBF as a unique stochastic algorithm parameterized by the memory function $(r_n)_{n \geq 1}$. From the definition of r_n given in (3.2.5), we note that in the exponential case, $r_n = r$ remains constant while the inertia brought by the memory term in the polynomial case $(r_n)_{n \in \mathbb{N}}$ is defined by $r_n = \frac{r}{\Gamma_n}$. Under Assumption $(\mathbf{H}_{\beta}^{\gamma})$, we can show that regardless of the memory, we have :

$$\sum_{n \in \mathbb{N}} \gamma_n r_n = +\infty.$$

This is true when $r_n = r$ because $\gamma_n = \gamma n^{-\beta}$ with $\beta \leq 1$. It is also true when we deal with a polynomial memory since in that case :

- if $\beta < 1$, then $\gamma_n r_n \sim \gamma n^{-\beta} \times r(1-\beta) \gamma^{-1} n^{-1+\beta} \sim r(1-\beta) n^{-1}$
- if $\beta = 1$, then $\gamma_n r_n \sim \frac{r}{n \log n}$ and $\sum_{k \leq n} \gamma_k r_k \sim \log(\log n)$.

Similarly, we also have that in the polynomial case, regardless of β :

$$\sum_n \gamma_n^2 r_n < +\infty,$$

although this bound holds in the exponential situation when $\beta > 1/2$. Below, we will use these properties on the sequences $(\gamma_n)_{n \geq 0}$ and $(r_n)_{n \geq 0}$ and define the next set of assumptions :

- Assumption $(\mathbf{H}_{\mathbf{r}})$: The sequence $(r_n)_{n \geq 0}$ is a non-increasing sequence such that :

$$\sum_{n \geq 1} \gamma_{n+1} r_n = +\infty \quad \text{and} \quad \sum_{n \geq 1} \gamma_{n+1}^2 r_n < +\infty \quad \text{and} \quad \limsup_{n \rightarrow +\infty} \frac{1}{2\gamma_{n+1}} \left(\frac{1}{r_n} - \frac{1}{r_{n-1}} \right) =: c_r < 1.$$

In the exponential case, $c_r = 0$, whereas if $r_n = r/\Gamma_n$, it can be shown that $c_r = \frac{1}{2r}$ and the last point is true when $r > 1/2$. In any case, r_∞ will refer to the limiting value of r_n when $n \rightarrow +\infty$, which is either 0 or $r > 0$.

3.2.4 Main results

Section 3.3 is dedicated to the situation of a general coercive function f . We obtain the almost sure convergence of the stochastic HBF towards a critical point of f .

Theorem 3.2.1. Assume that f satisfies (\mathbf{H}_s) , that $(\mathbf{H}_{\sigma,2})$ holds and that the sequences $(\gamma_n)_{n \geq 1}$ and $(r_n)_{n \geq 1}$ are chosen such that $(\mathbf{H}_\beta^\gamma)$ and (\mathbf{H}_r) are fulfilled. If for any z , $\{x, f(x) = z\} \cap \{x, \nabla f(x) = 0\}$ is locally finite, then (X_n) a.s. converges towards a critical point of f .

This result obviously implies the convergence when f has a unique critical point. In the next theorem, we focus on the case where this uniqueness assumption fails, under the additional elliptic assumption $(\mathbf{H}_\mathcal{E})$.

Theorem 3.2.2. Assume that f satisfies (\mathbf{H}_s) , that the noise is elliptic, *i.e.*, $(\mathbf{H}_\mathcal{E})$ holds, and the sequence $(\gamma_n)_{n \geq 1}$ is chosen such that $(\mathbf{H}_\beta^\gamma)$ and (\mathbf{H}_r) are fulfilled. If for any z , $\{x, f(x) = z\} \cap \{x, \nabla f(x) = 0\}$ is locally finite, we have :

- (a) If $r_n = r$ (exponential memory) and $(\mathbf{H}_{\sigma,2})$ holds, then (X_n) a.s. converges towards a local minimum of f .
- (b) If $r_n = r\Gamma_n^{-1}$ and the noise is sub-Gaussian, *i.e.*, $(\mathbf{H}_{Gauss,\sigma})$ holds, then (X_n) a.s. converges towards a local minimum of f when $\beta < 1/3$.

Remark 3.2.1. \triangleright The previous result provides some guarantees when f is a multiwell potential. In (a), we consider the exponentially weighted HBF and show that the convergence towards a local minimum of f always holds under the additional assumption $(\mathbf{H}_\mathcal{E})$. To derive this result, we will essentially use the former results of [33] on “homogeneous” stochastic algorithms.

\triangleright Point (b) is concerned by polynomially-weighted HBF and deserves more comment :

- First, the result is rather difficult because of the time inhomogeneity of the stochastic algorithm, which can be written as $Z_{n+1} = Z_n + \gamma_{n+1}F_n(Z_n) + \gamma_{n+1}\Delta M_{n+1}$: the drift term F_n depends on Z_n and on the integer n , which will induce technical difficulties in the proof of the result. In particular, the assumption $\beta < 1/3$ will be necessary to obtain a good lower bound of the drift term in the unstable manifold direction with the help of the Poincaré Lemma near hyperbolic equilibrium of a differential equation.
- Second, the sub-Gaussian assumption $(\mathbf{H}_{Gauss,\sigma})$ is less general than $(\mathbf{H}_{\sigma,2})$ even though it is still a reasonable assumption within the framework of a stochastic algorithm. To prove (b), we will need to control the fluctuations of the stochastic algorithm around its deterministic drift, which will be quantified by the expectation of the random variable $\sup_{k \geq n} \gamma_k^2 \|\Delta M_k\|^2$. The sub-Gaussian assumption will be mainly used to obtain an upper bound of such an expectation, with the help of a coupling argument. Our proof will follow a strategy used in [106] and [20] where this kind of expectation has to be upper bounded. Nevertheless, the novelty of our work is also to generalize the approach to unbounded martingale increments : the arguments of [106, 20] are only valid for a bounded martingale increment, which is a somewhat restrictive framework.

In Section 3.4, we focus on the consistency rate under stronger assumptions on the convexity of f . In the exponential memory case, we are able to control the quadratic error and to establish a CLT for the stochastic algorithm under the general assumption $(\mathbf{H}_{SC}(\alpha))$. In the polynomial case, the problem is more involved and we propose a result for the quadratic error only when f is

a quadratic function (see Remark 3.2.2 for further comments on this restriction). More precisely, using the notation \lesssim to refer to an inequality, up to a universal multiplicative constant, we establish the following results.

Theorem 3.2.3. Denote by x^* the unique minimizer of f and assume that $(\mathbf{H}_\beta^\gamma)$, (\mathbf{H}_s) , $(\mathbf{H}_{SC}(\alpha))$ and $(\mathbf{H}_{\sigma,2})$ hold, we have :

(a) When $r_n = r$ (exponential memory) and $\beta < 1$, we have :

$$\mathbb{E} [\|X_n - x^*\|^2 + \|Y_n\|^2] \lesssim \gamma_n$$

If $(\mathbf{H}_{\sigma,\infty})$ holds and $\beta = 1$, set $\alpha_r = r \left(1 - \sqrt{1 - \frac{(4\lambda)\wedge r}{r}}\right)$ where $\underline{\lambda}$ denotes the smallest eigenvalue of $D^2f(x^*)$. We have, for any $\varepsilon > 0$:

$$\mathbb{E} [\|X_n - x^*\|^2 + \|Y_n\|^2] \lesssim \begin{cases} n^{-1} & \text{if } \gamma\alpha_r > 1 \\ n^{-\alpha_r+\varepsilon} & \text{if } \gamma\alpha_r \leq 1. \end{cases}$$

(b) Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a quadratic function. Assume that $r_n = r\Gamma_n^{-1}$ (polynomial memory) with $\beta < 1$. Then, if $r > \frac{1+\beta}{2(1-\beta)}$, we have :

$$\mathbb{E} [\|X_n - x^*\|^2 + \Gamma_n \|Y_n\|^2] \lesssim \gamma_n$$

When $r_n = r\Gamma_n^{-1}$ (polynomial memory) and $\beta = 1$, we have :

$$\mathbb{E} [\|X_n - x^*\|^2 + \log n \|Y_n\|^2] \lesssim \frac{1}{\log n}.$$

For (a), the case $\beta < 1$ is a consequence of Proposition 3.4.3 (or Proposition 3.4.1 in the quadratic case), whereas the (more involved) case $\beta = 1$ is dealt with Propositions 3.4.1 and 3.4.4 for the quadratic and the non-quadratic cases, respectively. We first stress that that when $\beta < 1$, the noise only needs to satisfy $(\mathbf{H}_{\sigma,p})$ to obtain our upper bound. When we deal with $\beta = 1$, we could prove a positive result in the quadratic case when we only assume $(\mathbf{H}_{\sigma,p})$. Nevertheless, the stronger assumption $(\mathbf{H}_{\sigma,\infty})$ is necessary to produce a result in the general strongly convex situation. Finally, (b) is a consequence of Proposition 3.4.2.

Remark 3.2.2. \triangleright It is worth noting that in (a) ($\beta = 1$), the dependency of the parameter α_r in D^2f only appears through the smallest eigenvalue of $D^2f(x^*)$. In particular, it does not depend on $\inf_{x \in \mathbb{R}^d} \underline{\lambda}_{D^2f(x)}$ as it could be expected in this type of result. In other words, we are almost able to retrieve the conditions that appear when f is quadratic. This optimization of the constraint is achieved with a “power increase” argument, but this involves a stronger assumption $(\mathbf{H}_{\sigma,\infty})$ on the noise.

\triangleright The restriction to quadratic functions in the polynomial case may appear surprising. In fact, the “power increase” argument does not work in this non-homogeneous case. However, when $\beta < 1$, it would be possible to extend to non-quadratic functions through a Lyapunov argument (on this topic, see Remark 3.4.3), but under some quite involved conditions on r , β and the Hessian of f . Hence, we chose to only focus on the quadratic case and to try to obtain some potentially optimal conditions on r and β only (in particular, there is no dependence to the spectrum of D^2f). The interesting point is that it is possible to preserve the standard rate order when $\beta < 1$ but under the constraint $r > \frac{1+\beta}{2(1-\beta)}$, which increases with β . In particular, the rate $\mathcal{O}(n^{-1})$ cannot be attained

in this case (see Remark 3.4.2 for more details).

Finally, we conclude by a central limit theorem related to the stochastic algorithm the *exponential memory case*.

Theorem 3.2.4. Assume (\mathbf{H}_s) and $(\mathbf{H}_{SC}(\alpha))$ are true. Suppose that $r_n = r$ and that $(\mathbf{H}_\beta^\gamma)$ holds with $\beta \in (0, 1)$ or, $\beta = 1$ and $\gamma\alpha_r > 1$. Assume that $(\mathbf{H}_{\sigma,p})$ holds with $p > 2$ when $\beta < 1$ and $p = \infty$ when $\beta = 1$. Finally, suppose that the following condition is fulfilled :

$$\mathbb{E}[(\Delta M_{n+1})(\Delta M_{n+1})^t | \mathcal{F}_{n-1}] \xrightarrow{n \rightarrow +\infty} \mathcal{V} \quad \text{in probability} \quad (3.2.6)$$

where \mathcal{V} is a symmetric positive $d \times d$ -matrix. Let σ be a $d \times d$ -matrix such that $\sigma\sigma^t = \mathcal{V}$. Then,

- (i) The normalized algorithm $\left(\frac{X_n}{\sqrt{\gamma n}}, \frac{Y_n}{\sqrt{\gamma n}}\right)_n$ converges in law to a centered Gaussian distribution $\mu_\infty^{(\beta)}$, which is the invariant distribution of the (linear) diffusion with infinitesimal generator \mathcal{L} defined on \mathcal{C}^2 -functions by :

$$\mathcal{L}g(z) = \left\langle \nabla g(z), \left(\frac{1}{2\gamma} 1_{\{\beta=1\}} I_{2d} + H \right) z \right\rangle + \frac{1}{2} \text{Tr}(\Sigma^T D^2 g(z) \Sigma)$$

with

$$H = \begin{pmatrix} 0 & -I_d \\ rD^2 f(x^*) & -rI_d \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \sigma \end{pmatrix}.$$

- (ii) In the simple situation where $\mathcal{V} = \sigma_0^2 I_d$ ($\sigma_0 > 0$) and $\beta < 1$. In this case, the covariance of $\mu_\infty^{(\beta)}$ is given by

$$\frac{\sigma_0^2}{2} \begin{pmatrix} \{D^2 f(x^*)\}^{-1} & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & rI_d \end{pmatrix}$$

In particular,

$$\frac{X_n}{\sqrt{\gamma n}} \Rightarrow \mathcal{N}(0, \frac{\sigma_0^2}{2} \{D^2 f(x^*)\}^{-1}).$$

Remark 3.2.3. \triangleright As a first comment of the above theorem, let us note that in the fundamental example where :

$$\Delta M_{n+1} = \nabla f(X_n) - \partial_x F(X_n, \xi_n), \quad n \geq 1,$$

the additional assumption (3.2.6) is a continuity assumption. Actually, in this case :

$$\mathbb{E}[\Delta M_n \Delta M_n^t | \mathcal{F}_{n-1}] = \bar{\mathcal{V}}(X_n), \quad \text{with } \bar{\mathcal{V}}(x) = \text{Cov}(F(x, \xi_1)).$$

Thus, since $X_n \rightarrow x^*$ a.s., Assumption (3.2.6) is equivalent to the continuity of $\bar{\mathcal{V}}$ in x^* so that :

$$\mathcal{V} = \bar{\mathcal{V}}(x^*).$$

\triangleright Point (ii) of Theorem 3.2.4 reveals the behavior of the asymptotic variance of Y increases with r . This translates the fact that the instantaneous speed coordinate Y is proportional to r in Equation (3.2.3), which then implies a large variance of the Y coordinate when we use an important value of r .

\triangleright When $\beta = 1$, it is also possible (but rather technical) to make the limit variance explicit. The expression obtained with the classical stochastic gradient descent with step-size γn^{-1} and Hessian λ , the asymptotic variance is $\gamma/(2\lambda\gamma - 1)$, whose optimal value is attained when $\gamma = \lambda^{-1}$ (it attains

the Cramer-Rao lower bound). Concerning now the stochastic HBF, for example, when $d = 1$ and $r \geq 4\lambda$ (the result is still valid in higher dimensions, see Section 3.5), we can show that :

$$\lim_{n \rightarrow +\infty} \gamma_n^{-1} \mathbb{E}[X_n^2] = \sigma_0^2 \frac{2\lambda r \gamma^3}{(\gamma r - 1)(2\lambda\gamma - \check{\alpha}_-)(2\lambda\gamma - \check{\alpha}_+)},$$

where $\check{\alpha}_+ = 1 + \sqrt{1 - \frac{4\lambda}{r}}$ and $\check{\alpha}_- = 1 - \sqrt{1 - \frac{4\lambda}{r}}$. Similar expressions may be obtained when $r < 4\lambda$. Note also that we assumed that $\gamma\alpha_r > 1$, and it is easy to check that this condition implies that $\gamma r > 1$ because $\alpha_r \leq r$, regardless of r . In the meantime, this condition also implies that $2\lambda\gamma > \check{\alpha}_+ \geq \check{\alpha}_-$.

Finally, This explicit value could be used to find the optimal calibration of the parameters to obtain the best asymptotic variance. Unfortunately, the expressions are rather technical and we can see that such calibrations are far from being independent of λ , the a priori unknown Hessian of f on x^* .

3.3 Almost sure convergence of the stochastic heavy ball

In this section, the baseline assumption on the function f is (\mathbf{H}_s) , and we are thus interested in the almost sure convergence of the stochastic HBF. In particular, *we do not make any convexity assumption on f* .

Below, we will sometimes use standard and sometimes more intricate normalizations for the coupled process $Z_n = (X_n, Y_n)$. These normalizations will be of a different nature and, to be as clear as possible, we will always use the same notation Z_n and \check{Z}_n to refer to a rotation of the initial vector Z_n , whereas \tilde{Z}_n will introduce a scaling in the Y_n component of Z_n by a factor $\sqrt{r_n}$.

3.3.1 Preliminary result

We first state a useful upper bound that makes it possible to derive a Lyapunov-type control for the mean evolution of the stochastic algorithm $(X_n, Y_n)_{n \geq 1}$ described by (3.2.3). This result is based on the important function $(x, y) \mapsto V_n(x, y)$ that depends on two parameters $(a, b) \in \mathbb{R}_+^2$ defined by :

$$V_n(x, y) = (a + br_{n-1})f(x) + \frac{a}{2r_{n-1}}\|y\|^2 - b\langle \nabla f(x), y \rangle. \quad (3.3.7)$$

We will show that V_n plays the role of a (potentially time-dependent) Lyapunov function for the sequence $(X_n, Y_n)_{n \geq 1}$. The construction of V_n shares a lot of similarity with other Lyapunov functions built to control second-order systems. If the two first terms are classical and generate a $-\|y\|^2$ term, the last one is more specific to hypo-coercive dynamics and was already used in [62]. Recent works fruitfully exploit this kind of Lyapunov function (see, among others, the kinetic Fokker-Planck equations in [121] and the memory gradient diffusion in [53]). This function is obtained by the introduction of some Lie brackets of differential operators, leading to the presence of $\langle \nabla f(x), y \rangle$ that generates a mean reverting effect on the variable x .

Lemma 3.3.1. Assume that $(\mathbf{H}_{\sigma,2})$ and (\mathbf{H}_s) hold and suppose that $c_r < 1$. Then, for any $(a, b) \in \mathbb{R}_+^2$ such that :

$$\frac{a}{b} > \left(\frac{1}{2} \vee \frac{\|D^2 f\|_\infty}{1 - c_r} \vee r_\infty(c_f - 1) \right), \quad (3.3.8)$$

we have :

(i) A constant $C_1 > 0$ and an integer $n_0 \in \mathbb{N}$ exist such that for any $n \geq n_0$,

$$\forall x, y \in \mathbb{R}^d, \quad V_n(x, y) \geq C_1 \left(f(x) + \frac{\|y\|^2}{r_{n-1}} \right). \quad (3.3.9)$$

(ii) Some positive constants C_2, C_3 and $c_{a,b}$ exist such that :

$$\begin{aligned} \mathbb{E}[V_{n+1}(X_{n+1}, Y_{n+1}) | \mathcal{F}_n] \\ \leq V_n(X_n, Y_n)(1 + C_2 \gamma_{n+1}^2 r_n) - c_{a,b} \gamma_{n+1} \|Y_n\|^2 - b \gamma_{n+1} r_n \|\nabla f(X_n)\|^2 + C_3 \gamma_{n+1}^2 r_n. \end{aligned} \quad (3.3.10)$$

Proof :

Point (i) : For any non-negative u, v , the elementary inequality $uv \leq \frac{\rho}{2} u^2 + \frac{1}{2\rho} v^2$ holds for any $\rho > 0$. We apply this inequality with $u = \|\nabla f(x)\|$, $v = \|y\|$ and $\rho = 2r_n$ and obtain :

$$|\langle \nabla f(x), y \rangle| \leq r_{n-1} \|\nabla f(x)\|^2 + \frac{a}{4r_{n-1}} \|y\|^2.$$

It follows from Assumption (\mathbf{H}_s) that $\|\nabla f\|^2 \leq c_f f$. Using the above inequality, we obtain that for any $x, y \in \mathbb{R}^d$:

$$V_n(x, y) \geq (a + br_{n-1}(1 - c_f))f(x) + \frac{1}{2r_{n-1}} \left[a - \frac{b}{2} \right] \|y\|^2.$$

Choosing now a and b such that $a > b/2$ and $a > br_\infty(c_f - 1)$, we obtain the first assertion follows from (3.3.8). \diamond

Point (ii) : The Taylor formula ensures the existence of $\xi_{n+1,1}$ and $\xi_{n+1,2}$ in $[X_n, X_{n+1}]$ such that :

$$\begin{aligned} V_{n+1}(X_{n+1}, Y_{n+1}) &= (a + br_n) \left(f(X_n) - \gamma_{n+1} \langle \nabla f(X_n), Y_n \rangle + \frac{\gamma_{n+1}^2}{2} Y_n^t D^2 f(\xi_{n+1,1}) Y_n \right) \\ &+ \frac{a}{2r_n} (\|Y_n\|^2 + 2\gamma_{n+1} r_n (\langle Y_n, \nabla f(X_n) \rangle - \|Y_n\|^2 + \langle Y_n + \gamma_{n+1} r_n (\nabla f(X_n)) - Y_n, \Delta M_{n+1} \rangle) + \gamma_{n+1}^2 r_n^2 \|\Delta M_{n+1}\|^2) \\ &- b \langle \nabla f(X_n) - \gamma_{n+1} D^2 f(\xi_{n+1,2}) Y_n, Y_n + \gamma_{n+1} r_n (\nabla f(X_n) - Y_n + \Delta M_{n+1}) \rangle. \end{aligned}$$

Combining the similar terms leads to :

$$\begin{aligned} V_{n+1}(X_{n+1}, Y_{n+1}) &= V_n(X_n, Y_n) - b(r_n - r_{n-1})f(X_n) \\ &+ \gamma_{n+1} \langle \nabla f(X_n), Y_n \rangle \left(\underbrace{-a - br_n + a + br_n}_{=0} \right) - \gamma_{n+1} Y_n^t D_{n+1} Y_n - \gamma_{n+1} r_n b \|\nabla f(X_n)\|^2 \\ &+ \gamma_{n+1} r_n \Delta N_{n+1} + \gamma_{n+1} \Delta R_{n+1}, \end{aligned}$$

where $(\Delta N_n)_{n \geq 1}$ is a sequence of martingale increments, D_n is a $d \times d$ -matrix defined by :

$$D_{n+1} = a \left(1 - \frac{1}{2\gamma_{n+1}} \left(\frac{1}{r_n} - \frac{1}{r_{n-1}} \right) \right) I_d - b D^2 f(\xi_{n+1,2}),$$

and ΔR_{n+1} is a remainder term. Using (\mathbf{H}_s) , we know that D^2f is bounded, and we have the following bound for ΔR_{n+1} :

$$\|\Delta R_{n+1}\| \leq C_2 \gamma_{n+1} r_n (\|Y_n\|^2 + \|\Delta M_{n+1}\|^2 + \|\nabla f(X_n)\| \cdot \|Y_n\|),$$

where C_2 is a deterministic positive constant independent of n . The fact that $(r_n)_{n \geq 1}$ is a bounded sequence combined with Assumptions $(\mathbf{H}_{\sigma,2})$ and (\mathbf{H}_s) yields $\mathbb{E}[\|\Delta R_{n+1}\| | \mathcal{F}_n] \leq C_2 \gamma_{n+1} r_n (1 + \|Y_n\|^2 + f(X_n))$. It follows that :

$$\forall n \geq n_0 \quad \mathbb{E}[\|\Delta R_{n+1}\| | \mathcal{F}_n] \leq C_2 \gamma_{n+1} r_n V_n(X_n, Y_n).$$

Second, the condition given by (3.3.8) shows that an integer $n_1 \geq n_0$ and a constant $c_{a,b} > 0$ exist such that :

$$D_{n+1} Y_n^{\otimes 2} \geq c_{a,b} \|Y_n\|^2.$$

Using the previous bounds in $V_{n+1}(X_{n+1}, Y_{n+1})$ and the fact that $(r_n)_{n \in \mathbb{N}}$ is non-increasing shows that :

$$\exists n_2 \geq n_1 \quad \forall n \geq n_2 : \quad \mathbb{E}[V_{n+1}(X_{n+1}, Y_{n+1}) | \mathcal{F}_n] \leq V_n(X_n, Y_n) (1 + C \gamma_{n+1}^2 r_n) - c_{a,b} \gamma_{n+1} \|Y_n\|^2 - b \gamma_{n+1} r_n \|\nabla f(X_n)\|^2.$$

◇□

Note that if (\mathbf{H}_r) holds, then Equation (3.3.10) provides a strong repelling effect on the system (x, y) because in that case, $\sum \gamma_{n+1} r_n = +\infty$. This makes it possible to obtain a more precise a.s. convergence result, stated below.

Corollary 3.3.1. If $(\mathbf{H}_{\sigma,2})$ and (\mathbf{H}_s) hold and $(r_n)_{n \geq 1}$ satisfies (\mathbf{H}_r) , then we have :

(i)

$$\sup_{n \geq 1} \left(\mathbb{E}[f(X_n)] + \frac{1}{r_n} \mathbb{E}[\|Y_n\|^2] \right) < +\infty$$

(ii) $(V_n(X_n, Y_n))_{n \geq 1}$ is a.s.-convergent to $V_\infty \in \mathbb{R}_+$. In particular, $(X_n)_{n \geq 1}$ and $(Y_n/\sqrt{r_n})_{n \geq 1}$ are a.s.-bounded.

(iii) $\sum_{n \geq 1} \gamma_{n+1} r_n \left(\frac{\|Y_n\|^2}{r_n} + \|\nabla f(X_n)\|^2 \right) < +\infty$ a.s.

(iv) $(Y_n/\sqrt{r_n})_{n \geq 0}$ tends to 0 since $n \rightarrow +\infty$ and every limit point of $(X_n)_{n \geq 0}$ belong to $\{x, \nabla f(x) = 0\}$. Furthermore, if for any z , $\{x, f(x) = z\} \cap \{x, \nabla f(x) = 0\}$ is locally finite, $(X_n)_{n \geq 0}$ converges towards a critical point of f .

Proof of (i) – (ii) – (iii) : Under the conditions on (r_n) , we can check that some positive a and b exist such that the conclusions of the previous lemma hold true. We then deduce that :

$$\begin{aligned} \mathbb{E}[V_{n+1}(X_{n+1}, Y_{n+1}) | \mathcal{F}_n] \\ \leq V_n(X_n, Y_n) (1 + C \alpha_{n+1}) - U_{n+1}, \end{aligned}$$

with $\alpha_n = \gamma_n^2 r_n$ and $U_{n+1} = c_{a,b} \gamma_{n+1} \|Y_n\|^2 + b \gamma_{n+1} r_n \|\nabla f(X_n)\|^2$. Subsequently, using the Robbins-Siegmund Theorem (see, e.g., Theorem 3.7.1 in Section 3.7.1, borrowed from [44]), we deduce, on the one hand, that $\sup_{n \geq 1} \mathbb{E}[V_n(X_n, Y_n)] < +\infty$ and that $(V_n(X_n, Y_n))_{n \geq 1}$ almost surely (and in L^1) converge towards a random variable $V_\infty \in \mathbb{R}_+$. In particular, the coercivity of f implies the a.s.-boundedness of $(X_n)_{n \geq 0}$. On the other hand, the Robbins-Siegmund Theorem also implies that :

$$\sum_{n \geq 1} \gamma_{n+1} r_n \left(\frac{\|Y_n\|^2}{r_n} + \|\nabla f(X_n)\|^2 \right) < +\infty \quad a.s.$$

Hence, the three first statements follow. \diamond

Proof of (iv) : The proof relies on the so-called *ODE method* (see, e.g., [20]). Set $r_\infty = \lim_{n \rightarrow +\infty} r_n$. We deal with cases $r_\infty > 0$ and $r_\infty = 0$ separately.

Case $r_\infty > 0$ (exponential memory) : Set $\Gamma_n = \sum_{k=0}^n \gamma_k$ with the convention $\gamma_0 = 0$. Denote by $(\bar{z}(t))_{t \geq 0}$ the interpolated process defined by $\bar{z}(\Gamma_n) = Z_n = (X_n, Y_n)'$, $n \geq 0$, with linear interpolations between times Γ_n and Γ_{n+1} and let $\bar{z}^{(n)}$ be the associated *shifted-sequence* defined by :

$$\bar{z}^{(n)}(t) = \bar{z}(t + \Gamma_n) \quad t \geq 0.$$

Setting $\varepsilon_n = (0, (r_{n-1} - r_\infty)(\nabla f(X_n) - Y_n) + \Delta M_n)'$ and $h(x, y) = (-y, r_\infty(\nabla f(x) - y))'$, we have :

$$Z_{n+1} = Z_n + \gamma_{n+1}(h(Z_n) + \varepsilon_{n+1}).$$

Set $N(n, t) = \inf\{k \geq n, \gamma_{n+1} + \dots + \gamma_k \geq t\}$ (with the convention $\inf \emptyset = n$). Then, since $(Z_n)_{n \geq 0}$ is *a.s.*-bounded, it is a classical result on stochastic algorithm theory (see, e.g., [44], Theorem 9.2.8 and the remark below) that if for any $T > 0$,

$$\limsup_{n \rightarrow +\infty} \sup_{t \in [0, T]} \left\| \sum_{k=n+1}^{N(n, t)+1} \gamma_k \varepsilon_k \right\| = 0 \quad a.s., \quad (3.3.11)$$

then $(\bar{z}^{(n)})_{n \geq 0}$ is relatively compact (for the topology of uniform convergence on compact sets) and its limit points are solutions to the ODE $\dot{z} = h(z)$. Let us prove (3.3.11). Let $T > 0$. Using the Cauchy-Schwarz inequality, we have, for every $t \in [0, T]$:

$$\begin{aligned} \sum_{k=n+1}^{N(n, t)+1} \gamma_k (\|\nabla f(X_{k-1})\| + \|Y_k\|) &\leq \sqrt{2} \left(\sum_{k=n+1}^{N(n, t)+1} \gamma_k \right)^{\frac{1}{2}} \left(\sum_{k=n+1}^{N(n, t)+1} \gamma_k (\|\nabla f(X_{k-1})\|^2 + \|Y_{k-1}\|^2) \right)^{\frac{1}{2}} \\ &\leq \sqrt{2(T + \gamma_1)} \left(\sum_{k=n+1}^{+\infty} \gamma_k (\|\nabla f(X_{k-1})\|^2 + \|Y_{k-1}\|^2) \right)^{\frac{1}{2}} \xrightarrow{n \rightarrow +\infty} 0, \end{aligned} \quad (3.3.12)$$

where the last convergence follows from (iii). On the basis of Assumption $(\mathbf{H}_{\sigma, 2})$ and (iii), we also note that $(\langle \sum_{k=1}^n \gamma_k \Delta M_k \rangle)_{n \geq 1}$ is *a.s.*-convergent so that $\sum \gamma_n \Delta M_n$. It easily follows that :

$$\limsup_{n \rightarrow +\infty} \sup_{t \in [0, T]} \left\| \sum_{k=n+1}^{N(n, t)+1} \gamma_k \Delta M_k \right\| = 0 \quad a.s.$$

and that (3.3.11) is satisfied. Now, we again deduce from (3.3.12) that for any $T > 0$,

$$\sup_{t \leq T} \|\bar{z}^{(n)}(t) - \bar{z}^{(n)}(0)\| = \sup_{t \leq T} \|\bar{z}^{(n)}(t) - Z_n\| \xrightarrow{n \rightarrow +\infty} 0$$

so that each limit point is stationary. At this stage, we have thus proven that every limit point of $(\bar{z}^{(n)})_{n \geq 0}$ is a stationary solution to $\dot{z} = h(z)$. This implies that any limit point Z_∞ of $(Z_n)_{n \geq 0}$ satisfies $h(Z_\infty) = 0$ (and thus $Y_\infty = \nabla f(X_\infty) = 0$). Actually, let $(Z_{n_k})_{k \geq 1}$ be a convergent subsequence of the (*a.s.* bounded) sequence $(Z_n)_{n \geq 0}$ and denote its limit by Z_∞ . Up to a second

extraction, $(\bar{z}^{(n_k)})$ converges to a stationary solution \bar{z}^∞ of $\dot{z} = h(z)$. As a consequence, $h(\bar{z}^\infty(t)) = 0$ for any $t \geq 0$. In particular, $h(\bar{z}^\infty(0)) = h(Z_\infty) = 0$. By (ii) and the fact that $(Y_n)_{n \geq 0}$ converges to 0, we also deduce that $(f(X_n))_{n \geq 0}$ is *a.s.*-convergent. To conclude the proof, it remains to observe that the set of possible limits of subsequences of $(X_n)_{n \geq 1}$ is connected. This is true since $X_n - X_{n-1} = -\gamma_n Y_{n-1} \rightarrow 0$ as $n \rightarrow +\infty$. \diamond

Case $r_\infty = 0$ (polynomial memory) : In this case, the proof is somewhat similar but the identification of the asymptotic dynamics requires an appropriate normalization of Y_n ¹. Let us set :

$$\tilde{\gamma}_n = \gamma_n \sqrt{r_n}, \quad \tilde{\Gamma}_n = \sum_{k=0}^n \tilde{\gamma}_k, \quad \tilde{X}_n = X_n, \quad \tilde{Y}_n = \frac{Y_n}{\sqrt{r_n}}.$$

Also set by $\tilde{Z}_n = (\tilde{X}_n, \tilde{Y}_n)'$. The dynamic of \tilde{Z}_n is described by Lemma 3.3.2 below. We denote as $(\tilde{z}(t))_{t \geq 0}$ the interpolated process, *i.e.* defined by $\tilde{z}(\tilde{\Gamma}_n) = \tilde{Z}_n$, $n \geq 0$, with linear interpolations between times $\tilde{\Gamma}_n$ and $\tilde{\Gamma}_{n+1}$ and let $\tilde{z}^{(n)}$ be the associated *shifted-sequence* defined by

$$\tilde{z}^{(n)}(t) = \tilde{z}(t + \tilde{\Gamma}_n) \quad t \geq 0.$$

With this setting, the idea is to show that the sequence $(\tilde{z}^{(n)}(t))_{t \geq 0}$ is tight with limits being stationary solutions of a homogeneous O.D.E. $\dot{z} = \tilde{h}(z)$ (\tilde{h} being the drift to be determined). The sequence $(\tilde{Z}_n)_{n \geq 0}$ satisfies Lemma 3.3.2 that shows that $\tilde{Z}_{n+1} = \tilde{Z}_n + \tilde{\gamma}_{n+1} \left(\tilde{h}(\tilde{Z}_n) + \tilde{\varepsilon}_{n+1} \right)$ with $\tilde{h}(\tilde{x}, \tilde{y}) := (-\tilde{y}, \nabla f(\tilde{x}))'$ and :

$$\tilde{\varepsilon}_{n+1} = \begin{pmatrix} 0 \\ v_n^{(1)} \nabla f(\tilde{X}_n) + v_n^{(2)} \tilde{Y}_n + \sqrt{\frac{r_n}{r_{n+1}}} \Delta M_{n+1} \end{pmatrix},$$

where $v_n^{(1)}$ and $v_n^{(2)}$ are given in the statement of Lemma 3.3.2.

On the basis of Assumption (\mathbf{H}_r) , we know that : $\limsup_{n \rightarrow +\infty} \frac{1}{2\gamma_{n+1}} \left(\frac{1}{r_{n+1}} - \frac{1}{r_n} \right) < 1$ so that :

$$v_n^{(1)} = O\left(\frac{r_n - r_{n+1}}{r_{n+1}}\right) = O(\tilde{\gamma}_{n+1} \sqrt{r_n}) \quad \text{and} \quad v_n^{(2)} = O(\sqrt{r_n}).$$

Thus, $(v_n^{(1)})_{n \geq 1}$ and $(v_n^{(2)})_{n \geq 1}$ converge to 0 as $n \rightarrow +\infty$. We can now repeat the arguments used in the situation $r_\infty > 0$ and we obtain :

$$\limsup_{n \rightarrow +\infty} \sup_{t \in [0, T]} \left\| \sum_{k=n+1}^{\tilde{N}(n, t)+1} \tilde{\gamma}_k \tilde{\varepsilon}_k \right\| = 0 \quad a.s.,$$

where $\tilde{N}(n, t) = \inf\{k \geq n, \tilde{\gamma}_{n+1} + \dots + \tilde{\gamma}_k \geq t\}$. We can still combine (3.3.12) and (iii) to obtain $\sup_{t \leq T} |\tilde{z}^{(n)}(t) - \tilde{z}^{(n)}(0)| \xrightarrow{n \rightarrow +\infty} 0$ for any $T > 0$. We conclude that $(\tilde{z}^{(n)})_{n \geq 0}$ is relatively compact and that its limits are stationary solutions of $\dot{z} = \tilde{h}(z)$. The end of the proof is exactly the same as in the case $r_\infty > 0$. $\diamond \square$

1. In fact, due to the asymptotic stationarity, the limiting dynamics is not intrinsic.

Lemma 3.3.2.

$$\tilde{Z}_{n+1} = \tilde{Z}_n + \tilde{\gamma}_{n+1} \left(\tilde{h}(\tilde{Z}_n) + \tilde{\varepsilon}_{n+1} \right)$$

where $\tilde{h}(\tilde{x}, \tilde{y}) = (-\tilde{y}, \nabla f(\tilde{x}) - \sqrt{r_\infty} \tilde{y})'$ and

$$\tilde{\varepsilon}_{n+1} = \begin{pmatrix} 0 \\ v_n^{(1)} \nabla f(\tilde{X}_n) + v_n^{(2)} \tilde{Y}_n + \sqrt{\frac{r_n}{r_{n+1}}} \Delta M_{n+1} \end{pmatrix},$$

with

$$v_n^{(1)} = \sqrt{\frac{r_n}{r_{n+1}}} - 1 \quad \text{and} \quad v_n^{(2)} = \frac{1}{\gamma_{n+1}} v_n^{(1)} + \left(\sqrt{r_\infty} - \frac{r_n}{\sqrt{r_{n+1}}} \right).$$

Proof : First, the fact that $\tilde{X}_{n+1} = \tilde{X}_n - \tilde{\gamma}_{n+1} \tilde{Y}_n$ is obvious. Second,

$$\tilde{Y}_{n+1} = \tilde{Y}_n \sqrt{\frac{r_n}{r_{n+1}}} + \tilde{\gamma}_{n+1} \left(\sqrt{\frac{r_n}{r_{n+1}}} \nabla f(\tilde{X}_n) - \frac{r_n}{\sqrt{r_{n+1}}} \tilde{Y}_n + \sqrt{\frac{r_n}{r_{n+1}}} \Delta M_{n+1} \right).$$

The lemma follows. □

3.3.2 Convergence to a local minimum

To motivate the next theoretical result, we address the result of Corollary 3.3.1. We have shown the almost sure convergence of (3.2.3) towards a point of the form $(x_\infty, 0)$ in both exponential and polynomial cases where x_∞ is a critical point of f . This result is obtained under very weak assumptions on f and on the noise $(\Delta M_{n+1})_{n \geq 1}$ and is rather close to Theorems 3-4 of [122] (obtained within a different framework). Unfortunately, it this only provides a very partial answer to the problem of minimizing f because nothing is said about the stability of the limit of the sequence $(X_n)_{n \geq 0}$ by Corollary 3.3.1 : the attained critical point may be a local maximum, a saddle point or a local minimum.

This result is made more precise below and we establish some sufficient guarantees for the a.s. convergence of (X_n) towards a *minimum* of f , even if f possesses some local traps with the additional assumption $(\mathbf{H}_\mathcal{E})$. This proof follows the approach described in [33] and [20] but requires some careful adaptations because of the hypo-elliptic noise of the algorithm (there is no noise on the x -component) for both the exponentially and polynomially-weighted memory. Moreover, the linearization of the inhomogeneous drift around a critical point of f in the polynomial memory case is a supplementary difficulty we need to bypass.

Note that some recent works on stochastic algorithms (see, *e.g.*, [81]) deal with the convergence to minimizers of f of *deterministic* gradient descent with a randomized initialization. In our case, we will obtain a rather different result because of the randomization of the algorithm at each iteration. Note, however that the main ingredient of the proofs below will be the stable manifold theorem (the Poincaré Lemma on stable/unstable hyperbolic points of [108]) and its consequence around hyperbolic points. This geometrical result is also used in [81].

Exponential memory $r_n = r > 0$

The exponential memory case may be (almost) seen as an application of Theorem 1 of [33]. More precisely, if $Z_n = (X_n, Y_n)$ and $h(x, y) = (-y, r \nabla f(x) - ry)$, then the underlying stochastic algorithm may be written as :

$$Z_{n+1} = Z_n + \gamma_n h(Z_n) + \gamma_n \Delta M_n,$$

When $r_n = r > 0$ (exponential memory), Corollary 3.3.1 applies and $Z_n \xrightarrow{a.s.} Z_\infty = (X_\infty, 0)$ where X_∞ is a critical point of f . For the analysis of the dynamics around a critical point of the drift, the critical point of f is denoted x_0 and we can linearize the drift around $(x_0, 0) \in \mathbb{R}^d \times \mathbb{R}^d$ as :

$$h(x, y) = \begin{pmatrix} 0 & -I_d \\ rD^2(f)(x_0) & -rI_d \end{pmatrix} \begin{pmatrix} x - x_0 \\ y \end{pmatrix} + O(\|x - x_0\|^2),$$

where I_d is the $d \times d$ identity-squared matrix and $D^2(f)(x_0)$ is the Hessian matrix of f at point x_0 . When x_0 is not a local minimum of f , the spectral decomposition of $D^2(f)(x_0)$ leads to the spectral decomposition :

$$\exists P \in \mathcal{O}_d(\mathbb{R}) \quad D^2(f)(x_0) = P^{-1} \Lambda P,$$

where Λ is a diagonal matrix with at least one negative eigenvalue $\lambda < 0$. Considering now $Z_n = (X_n, Y_n)$ where $X_n = PX_n$ and $Y_n = PY_n$, we have :

$$Z_{n+1} = Z_n + \gamma_n \tilde{h}(Z_n) + \gamma_n P \Delta M_n,$$

where h may be linearized as :

$$h(x, y) = \begin{pmatrix} 0 & -I_d \\ r\Lambda & -rI_d \end{pmatrix} \begin{pmatrix} x - x_0 \\ y \end{pmatrix} + O(\|x - x_0\|^2) \quad \text{where} \quad x_0 = Px_0.$$

In particular, if e_λ is an eigenvector associated with the eigenvalue $\lambda < 0$ of $D^2f(x_0)$, we can see that the linearization of \tilde{h} on the space $\text{Span}(e_\lambda) \otimes (1, 0, \dots, 0)$ acts as :

$$A_{\lambda, r} = \begin{pmatrix} 0 & -1 \\ r\lambda & -r \end{pmatrix}.$$

Its spectrum is $Sp(A_{\lambda, r}) = -\frac{r}{2} \pm \sqrt{\frac{r^2}{4} - r\lambda}$. The important fact is that when $\lambda < 0$, the eigenvalue $-\frac{r}{2} + \sqrt{\frac{r^2}{4} - r\lambda}$ is positive and whose corresponding eigenspace is $E_\lambda^+ = \left(1, \frac{1}{2} - \sqrt{\frac{1}{4} - \lambda/r}\right)$. In the initial space $\mathbb{R}^d \times \mathbb{R}^d$ (without applying the change of basis through $P \otimes P$), the corresponding eigenvector is :

$$e_\lambda^+ = e_\lambda \otimes \left(\frac{1}{2} - \sqrt{\frac{1}{4} - \lambda/r}\right) e_\lambda$$

Consequently, when x_0 is not a local minimum of f , it generates a hyperbolic equilibrium of h and we can apply the “general” local trap Theorem 1 of [33]. If $\Pi_{E_\lambda^+}$ denotes the projection on the eigenspace $\text{Span}(e_\lambda^+)$, then the noise in the direction E_λ^+ is :

$$\xi_n^+ = \Pi_{E_\lambda^+}(0, \Delta M_n) = \frac{\langle \Delta M_n, e_\lambda \rangle}{\|e_\lambda\|^2} e_\lambda.$$

Now, Assumption $(\mathbf{H}_\mathcal{E})$ implies that :

$$\liminf_{n \rightarrow +\infty} \mathbb{E} \left\| \Pi_{E_\lambda^+}(0, \Delta M_n) \right\| \geq c_{e_\lambda} > 0.$$

We can then apply Theorem 1 of [33] and conclude the following result.

Theorem 3.3.1. If $(\mathbf{H}_{\sigma,2})$, (\mathbf{H}_s) and $(\mathbf{H}_{\mathcal{E}})$ hold and $r_n = r$, then X_n a.s. converges towards a local minimum of f .

Polynomial memory $r_n = r\Gamma_n^{-1} \rightarrow 0$

We introduce a key normalization of the speed coordinate and define the rescaled process :

$$\tilde{X}_n = X_n \quad \text{and} \quad \tilde{Y}_n = \sqrt{\Gamma_n} Y_n.$$

We can note that $\tilde{Y}_n = \sqrt{r} Y_n r_n^{-1/2}$ and the important conclusion brought by (iv) of Corollary 3.3.1 is that $(\tilde{X}_n, \tilde{Y}_n) \xrightarrow{a.s.} (X_\infty, 0)$ still holds (under the assumptions of Corollary 3.3.1). We can write the recursive upgrade of the couple $(\tilde{X}_n, \tilde{Y}_n)$. The evolution of $(\tilde{X}_n)_{n \geq 0}$ is easy to write : $\tilde{X}_{n+1} = \tilde{X}_n - \frac{\gamma_{n+1}}{\sqrt{\Gamma_n}} \tilde{Y}_n$. The recursive formula satisfied by $(\tilde{Y}_n)_{n \geq 0}$ is :

$$\begin{aligned} \tilde{Y}_{n+1} &= \sqrt{\Gamma_{n+1}} [Y_n + \gamma_{n+1} r_{n+1} (\nabla f(X_n) - Y_n + \Delta M_{n+1})] \\ &= \frac{\sqrt{\Gamma_{n+1}}}{\sqrt{\Gamma_n}} \tilde{Y}_n + r \frac{\gamma_{n+1}}{\sqrt{\Gamma_n}} \times \frac{\sqrt{\Gamma_{n+1}}}{\sqrt{\Gamma_n}} \nabla f(\tilde{X}_n) - r \frac{\gamma_{n+1}}{\sqrt{\Gamma_n}} \times \frac{\sqrt{\Gamma_{n+1}}}{\Gamma_n} \tilde{Y}_n + r \frac{\gamma_{n+1}}{\sqrt{\Gamma_n}} \times \frac{\sqrt{\Gamma_{n+1}}}{\sqrt{\Gamma_n}} \Delta M_{n+1} \end{aligned}$$

Hence, the couple $(\tilde{X}_n, \tilde{Y}_n)$ evolves as an almost standard stochastic algorithm, whose step size is $\tilde{\gamma}_{n+1} = \gamma_{n+1} \Gamma_n^{-1/2}$:

$$\begin{cases} \tilde{X}_{n+1} = \tilde{X}_n - \tilde{\gamma}_{n+1} \tilde{Y}_n \\ \tilde{Y}_{n+1} = \tilde{Y}_n + r \tilde{\gamma}_{n+1} \nabla f(\tilde{X}_n) + \tilde{\gamma}_{n+1} q_{n+1} \Delta M_{n+1} + \tilde{\gamma}_{n+1} U_{n+1}, \end{cases} \quad (3.3.13)$$

where $q_{n+1} = \sqrt{\Gamma_{n+1}/\Gamma_n} = 1 + o(n^{-1})$ as $n \rightarrow +\infty$ and $(U_{n+1})_{n \geq 1}$ is defined by :

$$U_{n+1} = \frac{1/2 - r q_{n+1} + o(n^{-1})}{\sqrt{\Gamma_n}} \tilde{Y}_n + r(q_{n+1} - 1) \nabla f(\tilde{X}_n).$$

This dynamical system is related to the deterministic one $\begin{cases} \dot{x}_t = -y_t \\ \dot{y}_t = r \nabla f(x_t) \end{cases}$ or equivalently :

$$\dot{z}_t = F(z_t) \quad \text{with} \quad F(z) = F(x, y) = (-y, r \nabla f(x)). \quad (3.3.14)$$

It is easy to see that when x_∞ is a local maximum of f , then the above drift is unstable near $z_\infty = (x_\infty, 0)$. Unfortunately, Theorem 1 of [33] cannot be applied because of the size of the remainder terms involved in (3.3.13) and the a.s. convergence of $(X_n, Y_n)_{n \geq 0}$ requires further investigation. From [20], we borrow a tractable construction of a ‘‘Lyapunov’’ function η in the neighborhood of each hyperbolic point, which translates a mean repelling effect of the unstable points. This construction still relies on the Poincaré Lemma (see [108] and [63] for a recent reference). Again, in the neighborhood of any hyperbolic point, we will treat the projection Π_+ as a projection on the unstable manifold.

Proposition 3.3.1 ([20]). For any local maximum point x_∞ of f , a compact neighborhood \mathcal{N} of $z_\infty = (x_\infty, 0)$ and a positive function $\eta \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}_+^d, \mathbb{R}_+^*)$ exist such that :

- (i) $\forall z = (x, y) \in \mathcal{N}, D\eta(z) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ is Lipschitz, convex and positively homogeneous.

(ii) Two constants $k > 0$ and $c_1 > 0$ and a neighborhood U of $(0, 0)$ exist such that :

$$\forall z \in \mathcal{N} \quad \forall u \in U \quad \eta(z + u) \geq \eta(z) + \langle D\eta(z), u \rangle - k\|u\|^2,$$

and if $\lfloor \cdot \rfloor_+$ denotes the positive part :

$$\forall z \in \mathcal{N} \quad \forall u \in U \quad \lfloor D\eta(z)(u) \rfloor_+ \geq c_1 \|\Pi_+(u)\|.$$

(iii) A positive constant κ exists such that :

$$\forall z \in \mathcal{N} \quad \langle D\eta(z), F(z) \rangle \geq \kappa \eta(z)$$

When $d = 1$, it is possible to check that if λ is a negative eigenvalue of the Hessian of f around a local maximum x_∞ , then the drift may be linearized in $(-y, \lambda(x - x_\infty))$ and a reasonable approximation of η is given by $\eta(x, y) = \frac{1}{2}\|y - \sqrt{-\lambda}x\|^2$. Nevertheless, the situation is more involved in higher dimensions and the construction of the function η relies on the stable manifold theorem. We are now able to state the next important result.

Theorem 3.3.2. Assume that the noise satisfies $(\mathbf{H}_{Gauss, \sigma})$ and $(\mathbf{H}_\mathcal{E})$, that the function satisfies (\mathbf{H}_s) , and that $\gamma_n = \gamma n^{-\beta}$ with $\beta < 1/3$, then $(X_n)_{n \geq 0}$ a.s. converges towards a local minimum of f .

The proof relies on an argument of [106, 20] even though it requires major modifications to deal with the time inhomogeneity of the process and the unbounded noise, which are assumed in these previous works. We denote \mathcal{N} as any neighborhood of z_∞ and consider any integer $n_0 \in \mathbb{N}$. We then introduce $\tilde{Z}_n = (\tilde{X}_n, \tilde{Y}_n)$ and the stopping time :

$$T := \inf \left\{ n \geq n_0 : \tilde{Z}_n \notin \mathcal{N} \right\}.$$

We will show that $\mathbb{P}(T < +\infty) = 1$, which implies the conclusion. We introduce two sequences $(\Omega_n)_{n \geq n_0}$ and $(S_n)_{n \geq n_0}$:

$$\Omega_{n+1} = [\eta(\tilde{Z}_{n+1}) - \eta(\tilde{Z}_n)] \mathbf{1}_{n < T} + \tilde{\gamma}_{n+1} \mathbf{1}_{n \geq T} \quad \text{and} \quad S_n = \eta(\tilde{Z}_{n_0}) + \sum_{k=n_0+1}^n \Omega_k. \quad (3.3.15)$$

Note that the construction of η implies that $z \mapsto D\eta(z)$ is Lipschitz, so that the following inequality holds :

$$\eta(z + u) - \eta(z) \geq \langle D\eta(z), u \rangle - \frac{\|D\eta\|_{Lip} \|u\|^2}{2}.$$

This inequality provides some information when u is small. In the meantime, η is positive so that :

$$\forall \alpha \in (0, 1] \quad \exists k_\alpha > 0 \quad \forall (z, u) \in \mathcal{N} \times \mathbb{R}^d \quad \eta(z + u) - \eta(z) \geq \langle D\eta(z), u \rangle - k_\alpha \|u\|^{1+\alpha} \quad (3.3.16)$$

The family of inequalities described in (3.3.16) will be used with an appropriate value of α in the next result.

Proposition 3.3.2. The random variables $(\Omega_n)_{n \geq 0}$ satisfy the following conditions :

(i) A constant c exists such that :

$$\mathbb{E}[\Omega_{n+1}^2 | \mathcal{F}_n] \leq c \tilde{\gamma}_{n+1}^2$$

(ii) A sequence $(\epsilon_n)_{n \geq 0}$ exists such that :

$$\mathbf{1}_{S_n \geq \epsilon_n} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \geq 0,$$

with $\epsilon_n \sim cn^{-(1-\alpha)/2}$ for a large enough c and $\alpha = (1 - \beta)/(1 + \beta)$.

(iii) Assume that $\beta < \frac{1}{3}$, then $(S_n^2)_{n \geq 0}$ has a submartingale increment :

$$\mathbb{E}[S_{n+1}^2 - S_n^2 | \mathcal{F}_n] \geq a\tilde{\gamma}_{n+1}^2$$

for a small enough constant a .

Proof :

Proof of (i). When $n \geq T$, we have $\Omega_{n+1} = \tilde{\gamma}_{n+1}$ by definition and the conclusion follows. In the other situation when $n \leq T$, we use the Lipschitz continuity of η : if $m = \sup_{z \in \mathcal{N}} \|D\eta(z)\|$, then Equation (3.3.13) yields :

$$\|\eta(\tilde{Z}_{n+1}) - \eta(\tilde{Z}_n)\|^2 \leq 4m^2\tilde{\gamma}_{n+1}^2 \left[\|\tilde{Y}_n\|^2 + r^2\|\nabla f(\tilde{X}_n)\|^2 + q_{n+1}^2\|\Delta M_{n+1}\|^2 + \|U_{n+1}\|^2 \right].$$

The neighborhood \mathcal{N} being compact, we deduce from the previous inequality that a constant $C > 0$ exists such that :

$$\mathbf{E} [\|\Omega_{n+1}\|^2 \mathbf{1}_{n < T} | \mathcal{F}_n] \leq \mathbf{E} [\|\eta(\tilde{Z}_{n+1}) - \eta(\tilde{Z}_n)\|^2 \mathbf{1}_{n < T} | \mathcal{F}_n] \leq C\tilde{\gamma}_{n+1}^2,$$

where we used a uniform upper bound on $\mathbb{E}[\|\Delta M_{n+1}\|^2 \mathbf{1}_{n < T} | \mathcal{F}_n]$, leading to the proof of (i). \diamond

Proof of (ii). Note that $\mathbf{1}_{n < T}$ and $\mathbf{1}_{n \geq T}$ are \mathcal{F}_n measurable and we have :

$$\mathbf{1}_{n \geq T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] = \mathbf{1}_{n \geq T} \tilde{\gamma}_{n+1} \geq 0.$$

On the complementary set, we also have :

$$\mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \geq \mathbf{1}_{n < T} \mathbb{E} [\eta(\tilde{Z}_{n+1}) - \eta(\tilde{Z}_n) | \mathcal{F}_n] = \mathbf{1}_{n < T} \mathbb{E} [\eta(\tilde{Z}_{n+1}) - \eta(\tilde{Z}_n) | \mathcal{F}_n]$$

Hence, we can use the lower bound given by (3.3.16) : for any value of $\alpha \in (0, 1]$:

$$\begin{aligned} \mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] &\geq \mathbf{1}_{n < T} \left[\tilde{\gamma}_{n+1} \langle D\eta(\tilde{Z}_n), F(\tilde{Z}_n) \rangle + \tilde{\gamma}_{n+1} \langle D\eta(\tilde{Z}_n), \mathbb{E}[\Delta M_{n+1} | \mathcal{F}_n] + U_{n+1} \rangle \right] \\ &\quad - \mathbf{1}_{n < T} k_\alpha \tilde{\gamma}_{n+1}^{1+\alpha} \left[\|\tilde{Y}_n\| + r\|\nabla f(\tilde{X}_n)\| + q_{n+1}\|\Delta M_{n+1}\| + \|U_{n+1}\| \right]^{1+\alpha} \end{aligned}$$

where we used the triangle inequality in the last line to derive an upper bound of $\|\tilde{Z}_{n+1} - \tilde{Z}_n\|$. When $n < T$, \tilde{Z}_n is bounded and we have $\mathbb{E}[\|\Delta M_{n+1}\|^2 | \mathcal{F}_n] \leq \sigma^2 M$ for a large enough M . Hence, the Hölder inequality implies that :

$$\mathbb{E}[\|\Delta M_{n+1}\|^{1+\alpha} | \mathcal{F}_n] \leq \sigma^{1+\alpha} M^{\frac{1+\alpha}{2}}.$$

Therefore, we can find a large enough constant $C_1 > 0$ such that :

$$\mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \geq \mathbf{1}_{n < T} \left[\tilde{\gamma}_{n+1} \langle D\eta(\tilde{Z}_n), F(\tilde{Z}_n) \rangle - m\tilde{\gamma}_{n+1}\|U_{n+1}\| - C_1\tilde{\gamma}_{n+1}^{1+\alpha} \right].$$

The lower bound (iii) of Proposition 3.3.1 and the definition of U_{n+1} implies that a constant C_2 exists such that :

$$\mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \geq \mathbf{1}_{n < T} \tilde{\gamma}_{n+1} \left[\kappa \eta(\tilde{Z}_n) - C_1 \tilde{\gamma}_{n+1}^\alpha - \frac{C_2}{\sqrt{\Gamma_n}} \right]$$

We now choose α so that $\tilde{\gamma}_{n+1}^\alpha \simeq \Gamma_n^{-1/2}$, which corresponds to the choice :

$$\alpha = \frac{1 - \beta}{1 + \beta}.$$

Defining $\epsilon_n = \kappa^{-1} \left[C_1 \tilde{\gamma}_{n+1} + C_2 \Gamma_n^{-1/2} \right]$, we then deduce that if $n < T$, then $S_n = \eta(\tilde{Z}_n)$ so that :

$$\mathbf{1}_{S_n \geq \epsilon_n} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \geq 0,$$

which concludes the proof. In particular, ϵ_n must be chosen on the order $\tilde{\gamma}_{n+1}^\alpha$ (or on the order $\Gamma_n^{-1/2} \sim n^{-(1-\beta)/2}$). \diamond

Proof of (iii). Observe that $S_{n+1}^2 - S_n^2 = \Omega_{n+1}^2 + 2S_n \Omega_{n+1}$. Now, if $S_n \geq \epsilon_n$, then we have seen in the proof of (ii) that :

$$\mathbf{1}_{S_n \geq \epsilon_n} \mathbb{E}[S_{n+1}^2 - S_n^2 | \mathcal{F}_n] = \mathbf{1}_{S_n \geq \epsilon_n} \mathbb{E}[\Omega_{n+1}^2 | \mathcal{F}_n] + 2S_n \mathbf{1}_{S_n \geq \epsilon_n} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \geq \mathbf{1}_{S_n \geq \epsilon_n} \mathbb{E}[\Omega_{n+1}^2 | \mathcal{F}_n].$$

In the other situation, we have $S_n \leq \epsilon_n$, meaning that $n < T$ and we have seen in the proof of (ii) that :

$$\begin{aligned} \mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] &\geq \mathbf{1}_{n < T} \left[\tilde{\gamma}_{n+1} \kappa \eta(\tilde{Z}_n) + \tilde{\gamma}_{n+1} \langle D\eta(\tilde{Z}_n), U_{n+1} \rangle \right] \\ &\quad - k_2 \tilde{\gamma}_{n+1}^2 \left[\|\tilde{Y}_n\| + r \|\nabla f(\tilde{X}_n)\| + q_{n+1} \|\Delta M_{n+1}\| + \|U_{n+1}\| \right]^2 \end{aligned}$$

Consequently, because of the positivity of η , we deduce that :

$$\mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \geq -\|D\eta(\tilde{Z}_n)\| \times O(\tilde{\gamma}_{n+1} \Gamma_n^{-1/2}) - O(\tilde{\gamma}_{n+1}^2).$$

We know that $D\eta$ is locally bounded on \mathcal{N} , we then obtain :

$$\begin{aligned} \mathbf{1}_{S_n \leq \epsilon_n} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] &= \mathbf{1}_{S_n \leq \epsilon_n} \mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] = \mathbf{1}_{\eta(\tilde{Z}_n) \leq \epsilon_n} \mathbf{1}_{n < T} \mathbb{E}[\Omega_{n+1} | \mathcal{F}_n] \\ &\geq -\mathbf{1}_{\eta(\tilde{Z}_n) \leq \epsilon_n} \mathbf{1}_{n < T} \left[\|D\eta(\tilde{Z}_n)\| \times O(\tilde{\gamma}_{n+1} \Gamma_n^{-1/2}) + O(\tilde{\gamma}_{n+1}^2) \right] \\ &\geq -C \tilde{\gamma}_{n+1} \left[\Gamma_n^{-1/2} + \tilde{\gamma}_{n+1} \right], \end{aligned}$$

for a large enough constant C . In the two situations, we then have :

$$\mathbb{E}[S_{n+1}^2 - S_n^2 | \mathcal{F}_n] \geq \mathbb{E}[\Omega_{n+1}^2 | \mathcal{F}_n] - 2C\epsilon_n \tilde{\gamma}_{n+1}^2 - 2C\epsilon_n \tilde{\gamma}_{n+1} \Gamma_n^{-1/2}.$$

Finally, Lemma 9.7 of [20] and our hypoelliptic assumption $(\mathbf{H}_{\mathcal{E}})$ implies that for small enough c :

$$\mathbb{E}[\Omega_{n+1}^2 | \mathcal{F}_n] \geq c \tilde{\gamma}_{n+1}^2$$

The conclusion follows if $\epsilon_n \tilde{\gamma}_{n+1} \Gamma_n^{-1/2} = o(\tilde{\gamma}_{n+1}^2)$. Since ϵ_n is chosen on the order $\Gamma_n^{-1/2} \sim \tilde{\gamma}_{n+1}^\alpha$ with $\alpha = (1 - \beta)/(1 + \beta)$, this condition is equivalent to :

$$\tilde{\gamma}_{n+1}^{1+2\alpha} = o(\tilde{\gamma}_{n+1}^2).$$

meaning that $\alpha > 1/2$. It then implies that β should be less than $1/3$. $\diamond \square$

We use now the key estimations derived from Proposition 3.3.2 to obtain the proof of the main result of this Section.

Proof of Theorem 3.3.2 : The proof is split into three parts. We consider :

$$S_n = S_0 + \sum_{k=1}^n \Omega_k \quad \text{and define} \quad \delta_n = \sum_{i \geq n} \tilde{\gamma}_i^2.$$

In our case, we have chosen $\beta \in (0, 1/3)$ and we can check that :

$$\tilde{\gamma}_n \sim n^{-(1+\beta)/2} \quad \text{so that} \quad \delta_n \sim n^{-\beta}. \quad (3.3.17)$$

We consider the sequence ϵ_n defined in Proposition 3.3.2 :

$$\epsilon_n \sim \Gamma_n^{-1/2} \sim \tilde{\gamma}_{n+1}^\alpha \quad \text{with} \quad \alpha = \frac{1-\beta}{1+\beta} > 1/2.$$

In this case, we have :

$$\epsilon_n = n^{-(1-\beta)/2} = o(n^{-\beta/2}) = o(\sqrt{\delta_n}) \quad \text{because} \quad \beta < 1/3 < 1/2.$$

The proof now proceeds by considering the sequential crossings $S_n \leq c\sqrt{\delta_n}$ and $S_n \geq c\sqrt{\delta_n}$ for a suitable value of c .

Step 1 : S_n becomes greater than $\sqrt{b\delta_n}$ with a positive probability.

For a given constant b and a positive $n \in \mathbb{N}$, we introduce the stopping time :

$$\mathcal{T} = \inf \left\{ i \geq n : S_i \geq \sqrt{b\delta_i} \right\},$$

and we show that an $\epsilon > 0$ exists such that $\mathbb{P}(\mathcal{T} < \infty) \geq 1 - \epsilon$. For a given by (iii) of Proposition 3.3.2, we consider :

$$\mathcal{M}_k = S_k^2 - a \sum_{i=0}^k \tilde{\gamma}_i^2.$$

$(\mathcal{M}_k)_{k \geq n}$ is a submartingale, so that $(\mathcal{M}_{k \wedge \mathcal{T}})_{k \geq n}$ is also a stopped submartingale. This yields :

$$\mathbb{E} [S_{m \wedge \mathcal{T}}^2 - S_n^2 | \mathcal{F}_n] \geq a \mathbb{E} \left[\sum_{i=n+1}^{m \wedge \mathcal{T}} \tilde{\gamma}_i^2 | \mathcal{F}_n \right] \geq a \left(\sum_{i=n+1}^m \tilde{\gamma}_i^2 \right) \mathbb{P}(\mathcal{T} > m | \mathcal{F}_n). \quad (3.3.18)$$

In the meantime, we can decompose $S_{m \wedge \mathcal{T}}^2 - S_n^2$ into :

$$\begin{aligned} S_{m \wedge \mathcal{T}}^2 - S_n^2 &= S_{m \wedge \mathcal{T}}^2 - S_{m \wedge \mathcal{T}-1}^2 + S_{m \wedge \mathcal{T}-1}^2 - S_n^2 \\ &\leq 2S_{m \wedge \mathcal{T}-1} \Omega_{m \wedge \mathcal{T}} + \Omega_{m \wedge \mathcal{T}}^2 + S_{m \wedge \mathcal{T}-1}^2 \\ &\leq 2S_{m \wedge \mathcal{T}-1}^2 + 2\Omega_{m \wedge \mathcal{T}}^2 \\ &\leq 2b\delta_{m \wedge \mathcal{T}-1} + 2\Omega_{m \wedge \mathcal{T}}^2. \end{aligned}$$

Since $(\delta_k)_{k \geq n}$ is decreasing, we then have $\delta_{m \wedge \mathcal{T}-1} \leq \delta_n$. We then study the remaining term. We can use Equation (3.3.13) and the Lipschitz continuity of η over the neighborhood \mathcal{N} (before time T) to obtain a large enough C such that :

$$\begin{aligned} \Omega_{m \wedge \mathcal{T}}^2 &= \Omega_{m \wedge \mathcal{T}}^2 [\mathbf{1}_{m \wedge \mathcal{T}-1 < T} + \mathbf{1}_{m \wedge \mathcal{T}-1 \geq T}] \\ &= [\eta(\tilde{Z}_{m \wedge \mathcal{T}}) - \eta(\tilde{Z}_{m \wedge \mathcal{T}-1})]^2 \mathbf{1}_{m \wedge \mathcal{T}-1 < T} + \tilde{\gamma}_{m \wedge \mathcal{T}}^2 \mathbf{1}_{m \wedge \mathcal{T}-1 \geq T} \\ &\leq C[\tilde{\gamma}_{m \wedge \mathcal{T}}^2 + \tilde{\gamma}_{m \wedge \mathcal{T}}^2 \|\Delta M_{m \wedge \mathcal{T}}\|^2]. \end{aligned}$$

However, nothing more is known about the stopped process $\|\Delta M_{m \wedge \mathcal{T}}\|^2$ and we are forced to use :

$$\mathbb{E} [S_{m \wedge \mathcal{T}}^2 - S_n^2 | \mathcal{F}_n] \leq 2b\delta_n + 2C \left[\tilde{\gamma}_n^2 + \mathbb{E} \left[\sup_{k \geq n} \tilde{\gamma}_k^2 \|\Delta M_k\|^2 \right] \right].$$

Given that all ΔM_k are independent sub-Gaussian random variables that satisfy Inequality (3.7.56), we can use Theorem 3.7.2 and obtain that a constant C large enough exists such that for any $\epsilon > 0$:

$$\mathbb{E} [S_{m \wedge \mathcal{T}}^2 - S_n^2 | \mathcal{F}_n] \leq 2b\delta_n + 2C\tilde{\gamma}_n^2 \log(\tilde{\gamma}_n^{-2}). \quad (3.3.19)$$

We can plug the estimate (3.3.19) into Inequality (3.3.18) to obtain :

$$\mathbb{P}(\mathcal{T} > m | \mathcal{F}_n) \leq \frac{2b\delta_n + 2C\tilde{\gamma}_n^2 \log(\tilde{\gamma}_n^{-2})}{a \sum_{i=n+1}^m \tilde{\gamma}_i^2}.$$

Letting $m \rightarrow +\infty$, we deduce that :

$$\mathbb{P}(\mathcal{T} = \infty | \mathcal{F}_n) \leq \frac{2b}{a} + \frac{2C\tilde{\gamma}_n^2 \log(\tilde{\gamma}_n^{-2})}{a\delta_n}.$$

According to the calibration (3.3.17), we have $\tilde{\gamma}_n^2 \log(\tilde{\gamma}_n^{-2}) = o(\delta_n)$. Consequently, we can choose n large enough such that :

$$\mathbb{P}(\mathcal{T} < \infty | \mathcal{F}_n) \geq 1 - \frac{3b}{a}.$$

◇

Step 2 : The sequence $(S_k)_{k \geq n}$ may remain larger than $\sqrt{b/2\delta_n}$ with a positive probability.

We introduce the stopping time \mathcal{S} and the event $E_n \in \mathcal{F}_n$:

$$\mathcal{S} = \inf\{i \geq n : S_i < \frac{\sqrt{b}}{2} \sqrt{\delta_n}\} \quad \text{and} \quad E_n = \{S_n \geq \sqrt{b} \sqrt{\delta_n}\}.$$

Since the sequence $(\delta_i)_{i \geq n}$ is non-increasing, (ii) of Proposition 3.3.2 yields :

$$\begin{aligned}\mathbb{E}[S_{(i+1) \wedge \mathcal{S}} - S_{i \wedge \mathcal{S}} | \mathcal{F}_i] &= \mathbf{1}_{\mathcal{S} > i} \mathbb{E}[S_{i+1} - S_i | \mathcal{F}_i] = \mathbf{1}_{\mathcal{S} > i} \mathbf{1}_{S_i \geq \sqrt{b/2\delta_n}} \mathbb{E}[S_{i+1} - S_i | \mathcal{F}_i] \\ &\geq \mathbf{1}_{\mathcal{S} \geq i} \mathbf{1}_{S_i \geq \sqrt{b/2\delta_i}} \mathbb{E}[X_{i+1} | \mathcal{F}_i] \geq \mathbf{1}_{\mathcal{S} \geq i} \mathbf{1}_{S_i \geq \epsilon_i} \mathbb{E}[X_{i+1} | \mathcal{F}_i] \geq 0.\end{aligned}$$

Hence, $(S_{i \wedge \mathcal{S}})_{i \geq n}$ is a submartingale and the Doob decomposition reads $S_{i \wedge \mathcal{S}} = M_i + I_i$ where $(M_i)_{i \geq n}$ is a Martingale and (I_i) is a predictable increasing process such that $I_n = 0$. Hence,

$$\mathbb{P}(\mathcal{S} = \infty | \mathcal{F}_n) = \mathbb{P}_{|\mathcal{F}_n} \left(\forall i \geq n : S_i \geq \frac{\sqrt{b}}{2} \sqrt{\delta_n} \right) \geq \mathbb{P}_{|\mathcal{F}_n} \left(\forall i \geq n : M_i \geq \frac{\sqrt{b}}{2} \sqrt{\delta_n} \right)$$

On the event E_n , $S_n = M_n \geq \sqrt{b}\sqrt{\delta_n}$ so that $M_i - M_n \leq M_i - \sqrt{b}\sqrt{\delta_n}$. Therefore :

$$\mathbb{P} \left(\forall i \geq n : M_i \geq \frac{\sqrt{b}}{2} \sqrt{\delta_n} | \mathcal{F}_n \right) \mathbf{1}_{E_n} \geq \mathbb{P} \left(\forall i \geq n : M_i - M_n \geq -\frac{\sqrt{b}}{2} \sqrt{\delta_n} | \mathcal{F}_n \right) \mathbf{1}_{E_n}.$$

The rest of the proof follows a standard martingale argument :

$$\begin{aligned}\mathbb{E}((M_i - M_n)^2 | \mathcal{F}_n) &= \sum_{j=n}^{i-1} \mathbb{E}((M_{j+1} - M_j)^2 | \mathcal{F}_n) = \sum_{j=n}^{i-1} \mathbb{E}(\mathbb{E}((M_{j+1} - M_j)^2 | \mathcal{F}_j) | \mathcal{F}_n) \\ &= \sum_{j=n}^{i-1} \mathbb{E}(\mathbb{E}((S_{j+1} - S_j)^2 | \mathcal{F}_j) - (I_{j+1} - I_j)^2 | \mathcal{F}_n) \\ &\leq \sum_{j=n}^{i-1} \mathbb{E}((S_{j+1} - S_j)^2 | \mathcal{F}_n) \leq \sum_{j=n}^{i-1} \mathbb{E}(\Omega_{j+1}^2 | \mathcal{F}_n) \leq c \sum_{j=n}^i \tilde{\gamma}_{j+1}^2 \leq c\delta_n.\end{aligned}$$

where we used the upper bound given by (i) of Proposition 3.3.2 in the last line. Now, the Doob inequality implies that :

$$\begin{aligned}\mathbb{P}(\inf_{n \leq i \leq m} (M_i - M_n) \leq -s | \mathcal{F}_n) &= \mathbb{P}(\inf_{n \leq i \leq m} (M_i - M_n - t) \leq -s - t | \mathcal{F}_n) \\ &\leq \mathbb{P}(\sup_{n \leq i \leq m} |M_i - M_n - t| \leq s + t | \mathcal{F}_n) \\ &\leq \frac{\mathbb{E}((M_m - M_n - t)^2 | \mathcal{F}_n)}{(s + t)^2} \\ &= \frac{\mathbb{E}((M_m - M_n)^2 | \mathcal{F}_n) + t^2}{(s + t)^2} = \frac{c\delta_n + t^2}{(s + t)^2}.\end{aligned}$$

We apply this inequality with $s = \frac{\sqrt{b}}{2} \sqrt{\delta_n}$ and use $(s + t)^2 \leq (1 + \vartheta)s^2 + (1 + \vartheta^{-1})t^2$ for any $\vartheta > 0$. It leads to :

$$\mathbb{P} \left(\inf_{n \leq i \leq m} (M_i - M_n) \leq -\frac{\sqrt{b}}{2} \sqrt{\delta_n} | \mathcal{F}_n \right) \leq \frac{c\delta_n + t^2}{(1 + \vartheta)b\delta_n/4 + (1 + \vartheta^{-1})t^2}.$$

We now choose $\vartheta = 4c/b$, $t = \sqrt{\delta_n}$ and deduce that :

$$\mathbb{P} \left(\inf_{n \leq i \leq m} (M_i - M_n) \leq -\frac{\sqrt{b}}{2} \sqrt{\delta_n} | \mathcal{F}_n \right) \leq \frac{c+1}{c+1+b/4c}.$$

Consequently, we deduce that :

$$\mathbb{P}(S = \infty | \mathcal{F}_n) \mathbf{1}_{E_n} \geq \mathbb{P}_{|\mathcal{F}_n} \left(\forall i \geq n : M_i \geq \frac{\sqrt{b}}{2} \sqrt{\delta_n} \right) \mathbf{1}_{E_n} \geq \left(1 - \frac{c+1}{c+1+b/4c} \right) \mathbf{1}_{E_n} = \frac{b}{b+4c+4c^2} \mathbf{1}_{E_n}$$

◇

Step 3 : $(S_n)_{n \geq 0}$ does not converge to 0 with probability 1.

We denote \mathcal{G} as the event that $(S_n)_{n \geq 0}$ does not converge to 0. For any integer n , we have the inclusion :

$$\{S = +\infty\} = \left\{ \forall i \geq n : S_i \geq \sqrt{b/4} \sqrt{\delta_n} \right\} \subset \mathcal{G},$$

which implies :

$$\mathbb{E}[\mathbf{1}_{\mathcal{G}} | \mathcal{F}_i] \mathbf{1}_{\mathcal{T}=i} = \mathbb{E}[\mathbf{1}_{\mathcal{G}} | \mathcal{F}_i] \mathbf{1}_{\mathcal{T}=i} \mathbf{1}_{E_i} \geq \frac{b}{b+4c+4c^2} \mathbf{1}_{\mathcal{T}=i} \mathbf{1}_{E_i} = \frac{b}{b+4c+4c^2} \mathbf{1}_{\mathcal{T}=i}$$

Hence,

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{\mathcal{G}} | \mathcal{F}_n] &= \sum_{i \geq n} \mathbb{E}[\mathbf{1}_{\mathcal{G}} \mathbf{1}_{\mathcal{T}=i} | \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[\mathbf{1}_{\mathcal{G}} | \mathcal{F}_i] \mathbf{1}_{\mathcal{T}=i} | \mathcal{F}_n] \\ &\geq \frac{b}{b+4c+4c^2} \sum_{i \geq n} \mathbb{E}[\mathbf{1}_{\mathcal{T}=i} | \mathcal{F}_n] \geq \frac{b}{b+4c+4c^2} \mathbb{P}(\mathcal{T} < +\infty | \mathcal{F}_n) \geq \frac{b}{b+4c+4c^2} \left(1 - \frac{3b}{a} \right) > 0. \end{aligned}$$

Since $\mathbf{1}_{\mathcal{G}} \in \mathcal{F}_{\infty}$, we have $\lim_{n \rightarrow +\infty} \mathbb{E}[\mathbf{1}_{\mathcal{G}} | \mathcal{F}_n] = \mathbf{1}_{\mathcal{G}}$. The previous lower bound implies that \mathcal{G} almost surely holds. ◇

Conclusion of the proof : *The stochastic algorithm does not converge to a local trap.*

Consider \mathcal{N} a neighborhood of a local maximum of f , and its associated function η given by Proposition 3.3.1. We then consider the random variables $(\Omega_n)_{n \geq 0}$ and $(S_n)_{n \geq 0}$. We have seen that S_n does not converge to 0 with probability 1. We define :

$$\mathcal{T}_{\mathcal{N}} := \inf \left\{ n \geq 0 : \tilde{Z}_n \notin \mathcal{N} \right\}.$$

and assume that $\mathcal{T}_{\mathcal{N}} = +\infty$. In that case, we always have :

$$\Omega_{n+1} = \eta(\tilde{Z}_{n+1}) - \eta(\tilde{Z}_n) \quad \text{and} \quad S_n = \eta(\tilde{Z}_n).$$

The limit set of $(\tilde{Z}_n)_{n \geq 0}$ is a non empty compact subset of $\overline{\mathcal{N}}$, which is left invariant by the flow $(\Phi_t)_{t \geq 0}$ of the O.D.E. whose drift is F . Now, consider z in $(\tilde{Z}_n)_{n \geq 0}$ and apply (iii) of Proposition 3.3.1. We then have $\eta(\Phi_t(z)) \geq e^{\kappa t} \eta(y)$. Since $\eta(\Phi_t(z)) \leq \sup_{\mathcal{N}} \eta$, we therefore deduce that $\eta(z) = 0$. Hence, the unique limiting value for $(S_n)_{n \geq 0}$ is zero, meaning that $S_n \rightarrow 0$ as $n \rightarrow +\infty$. However, we have seen in Step 3 that S_n does not converge to 0 with probability 1. Therefore, $\mathbb{P}(\mathcal{T}_{\mathcal{N}} = +\infty) = 0$ and the process does not converge towards a local maximum of f with probability

1. □

3.4 Convergence rates for strongly convex functions

This section focuses on the convergence rates of algorithm (3.2.3) according to the step-size $\gamma_n = \gamma n^{-\beta}$ for λ -strongly convex function f with a L -Lipschitz gradient, corresponding to the assumptions $(\mathbf{H}_{SC}(\lambda))$ and (\mathbf{H}_s) .

3.4.1 Quadratic case

We first study the benchmark case of a purely quadratic function f , meaning that ∇f is linear. In this case, $f(x) = \frac{1}{2}\|Ax\|^2$ and $\nabla f(x) = Sx$, leading to the following form of the algorithm :

$$\begin{cases} X_{n+1} = X_n - \gamma_{n+1}Y_n \\ Y_{n+1} = Y_n + \gamma_{n+1}r_n(SX_n - Y_n) + \gamma_{n+1}r_n\Delta M_{n+1}, \end{cases} \quad (3.4.20)$$

where S is a $d \times d$ squared matrix defined by $S = A'A$. The matrix S is assumed to be positive definite with lower bounded eigenvalues, *e.g.*, $Sp(S) \subset [\lambda, +\infty[$ when f is $(\mathbf{H}_{SC}(\lambda))$ with $\lambda > 0$.

Reduction to a two dimensional system

Equation (3.4.20) may be parameterized in a simpler form using the spectral decomposition of $S = P^{-1}\Lambda P$, where P is orthogonal, and Λ is a diagonal matrix :

$$\forall (i, j) \in \{1 \dots d\}^2 \quad \Lambda_{i,j} = \lambda_i \delta_{i,j} \geq \lambda > 0.$$

Keeping the notation $(X_n, Y_n)_{n \geq 1}$ for the change of basis induced by P , we define $X_n = PX_n$ and $Y_n = PY_n$ and obtain :

$$\begin{cases} X_{n+1} = X_n - \gamma_{n+1}\tilde{Y}_n \\ Y_{n+1} = Y_n + \gamma_{n+1}r_n(\Lambda X_n - Y_n) + \gamma_{n+1}r_n P \Delta M_{n+1}, \end{cases}$$

Since Λ is diagonal, we are now led to study the evolution of d couples of stochastic algorithms :

$$\forall i \in \{1 \dots d\} \quad \begin{cases} x_{n+1}^{(i)} = x_n^{(i)} - \gamma_{n+1}y_n^{(i)} \\ y_{n+1}^{(i)} = y_n^{(i)} + \gamma_{n+1}r_n(\lambda_i x_n^{(i)} - y_n^{(i)}) + \gamma_{n+1}r_n \Delta M_{n+1}^{(i)}, \end{cases}$$

where we used the notations $X_n = (x_n^{(i)})_{1 \leq i \leq d}$ and $Y_n = (y_n^{(i)})_{1 \leq i \leq d}$. Consequently, in the quadratic case, the stochastic HBF may be reduced to d couples of 2-dimensional random dynamical systems :

$$\forall i \in \{1, \dots, d\}^2 \quad Z_{n+1}^{(i)} = (I_2 + \gamma_{n+1}C_n^{(i)})Z_n^{(i)} + \gamma_{n+1}r_n \Sigma_2 \Delta N_{n+1}^{(i)}, \quad (3.4.21)$$

where

$$Z_n^{(i)} := (x_n^{(i)}, y_n^{(i)}) \quad \text{and} \quad C_n^{(i)} = \begin{pmatrix} 0 & -1 \\ \lambda^{(i)} r_n & -r_n \end{pmatrix} \quad \text{and} \quad \Sigma_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$\lambda^{(i)} = \Lambda_{i,i} \geq \lambda > 0$ and $(\Delta N_n^{(i)})_{n \geq 1}$ is a sequence of martingale increments.

It is worth noting that due to the multiplication by the matrix P , the martingale increment $\Delta N_{n+1}^{(i)}$ potentially depends on the whole coordinate $(Z_n^{(j)})_{1 \leq j \leq d}$. In a completely general case, this involves technicalities mainly due to the fact that the system (3.4.21) is not completely autonomous (in general, the components $Z_n^{(i)}$ and $Z_n^{(j)}$ do not evolve independently). To overcome this difficulty, the idea is to obtain some general controls for a system solution to (3.4.21) and to then bring the controls of each coordinate together. For the sake of simplicity, we propose in the sequel to state the results in the general case but to only make the proof for (3.4.21) with the assumption that :

$$\mathbb{E}[|\Delta N_{n+1}^{(j)}|^2 | \mathcal{F}_n] \leq C(1 + \|X_n^{(j)}\|^2). \quad (3.4.22)$$

From now on, we will omit the indexation by j to alleviate the notations. An easy computation shows that the characteristic polynomial of C_n is given by :

$$\chi_{C_n}(t) = \left(t + \frac{r_n}{2}\right)^2 + \frac{r_n(4\lambda - r_n)}{4}.$$

We now consider the two different cases :

- For all $n \geq 1$, C_n has two real or complex eigenvalues whose values do not change from n to n , which corresponds to $r_n = r$. This case necessarily corresponds to an exponentially-weighted memory and r_n is thus kept fixed constant : $r_n = r \geq 4\lambda$ or $r_n = r < 4\lambda$.
- For a large enough n , C_n has two complex conjugate and vanishing eigenvalues. This situation may occur if we use a polynomially-weighted memory because, in that case, $r_n \rightarrow 0$ as $n \rightarrow +\infty$.

Exponential memory $r_n = r$

We first study the situation when $r_n = r$, which is easier to deal with from a technical point of view.

Proposition 3.4.1. Assume $(H_{\sigma,2})$. Let $(Z_n)_{n \geq 0}$ be defined by (3.4.20) with $Sp(S) \subset [\lambda, +\infty[$ and $r_n = r$. Set :

$$\alpha_r = \begin{cases} r \left(1 - \sqrt{1 - \frac{4\lambda}{r}}\right), & \text{if } r \geq 4\lambda \\ r & \text{if } r < 4\lambda, \end{cases}.$$

Assume that $\gamma_n = \gamma n^{-\beta}$, we then have :

- (i) If $\beta < 1$, then a constant $c_{r,\lambda,\gamma}$ exists such that :

$$\forall n \geq 1 \quad \mathbb{E} [\|X_n\|^2 + \|Y_n\|^2] \leq c_{r,\lambda,\gamma} \gamma_n.$$

- (ii) If $\beta = 1$, then a constant $c_{r,\lambda,\gamma}$ exists such that :

$$\forall n \geq 1 \quad \mathbb{E} [\|X_n\|^2 + \|Y_n\|^2] \leq c_{r,\lambda,\gamma} n^{-(1 \wedge \gamma \alpha_r)} \log(n) \mathbf{1}_{\{\gamma \alpha_r = 1\}}.$$

Proof : According to Subsection 3.4.1, we only make the proof for a system solution to (3.4.21) with the assumption that (3.4.22) holds. We begin with the simplest case where $r \geq 4\lambda$. The above

computations show that :

$$Sp(C_n) = \left\{ \mu_+ = \frac{-r + \sqrt{(r-4\lambda)r}}{2}; \mu_- = \frac{-r - \sqrt{(r-4\lambda)r}}{2} \right\}, \quad (3.4.23)$$

while the associated eigenvectors are given by $e_+ = \begin{pmatrix} 1 \\ -\mu_+ \end{pmatrix}$ and $e_- = \begin{pmatrix} 1 \\ -\mu_- \end{pmatrix}$ and are kept fixed throughout the iterations of the algorithm. Consequently, (3.4.21) may be rewritten in an even simpler way :

$$Z_{n+1} = \begin{pmatrix} 1 + \gamma_{n+1}\mu_+ & 0 \\ 0 & 1 + \gamma_{n+1}\mu_- \end{pmatrix} Z_n + r\gamma_{n+1}\xi_{n+1}, \quad (3.4.24)$$

where $Z_n = QZ_n$ ((Z_n) being defined by (3.4.21)) where Q is an invertible matrix such that $C_n = Q^{-1} \begin{pmatrix} \mu_+ & 0 \\ 0 & \mu_- \end{pmatrix} Q$ and $\xi_{n+1} = Q\Sigma_2\Delta N_{n+1}$. The squared norm of $(Z_n)_{n \geq 1}$ is now controlled using a standard martingale argument and Assumption $(\mathbf{H}_{\sigma,2})$:

$$\mathbb{E} [\|Z_{n+1}\|^2 | \mathcal{F}_n] \leq [(1 + \mu_+\gamma_{n+1})^2 + C_1\gamma_{n+1}^2] \|Z_n\|^2 + C_2\gamma_{n+1}^2,$$

so that by setting $u_n = \mathbb{E}[\|Z_n\|^2]$, this yields :

$$u_{n+1} \leq (1 + 2\mu_+\gamma_{n+1} + C_1\gamma_{n+1}^2) + C_2\gamma_{n+1}^2. \quad (3.4.25)$$

The result then follows from Propositions 3.7.1 (iii) and 3.7.2 (iii) (see Appendix 3.7).

We now study the situation $r < 4\lambda$. In this case, C_n possesses two conjugate complex eigenvalues :

$$Sp(C_n) = \left\{ \mu_+ = \frac{-r + i\sqrt{r(4\lambda - r)}}{2}; \mu_- = \frac{-r - i\sqrt{r(4\lambda - r)}}{2} \right\},$$

Once again, we use the notation $(Z_n)_{n \geq 1}$ defined as $Z_n = QZ_n$ with Q an invertible (complex) matrix such that $S_n = Q^{-1} \begin{pmatrix} \mu_+ & 0 \\ 0 & \mu_- \end{pmatrix} Q$ and $\xi_{n+1} = Q\Sigma_2\Delta N_{n+1}$. The squared norm of $(Z_n)_{n \geq 1}$ may be controlled while paying attention to the modulus of complex numbers, and we obtain an inequality similar to (3.4.25).

$$\begin{aligned} \mathbb{E} [\|Z_{n+1}\|^2 | \mathcal{F}_n] &\leq \max \left(|1 + \mu_+\gamma_{n+1}|^2; |1 + \mu_-\gamma_{n+1}|^2 \right) \|Z_n\|^2 + C_2\gamma_{n+1}^2, \\ &\leq \left(\left(1 - \frac{\gamma_{n+1}r}{2} \right)^2 + C_1\gamma_{n+1}^2 \right) \|Z_n\|^2 + C_2\gamma_{n+1}^2, \\ &\leq (1 - \gamma_{n+1}r + C_1\gamma_{n+1}^2) \|Z_n\|^2 + C_2\gamma_{n+1}^2. \end{aligned}$$

Once again, we can apply (iii) of Propositions 3.7.1(iii) and 3.7.2(iii) to obtain the desired conclusion. \square

Remark 3.4.1. In the above proposition, the constants $c_{r,\lambda,\gamma}$ are not made explicit. However, it is possible to obtain an estimation if we assume that $\mathbb{E}[\Delta M_{n+1}|^2] \leq \sigma^2$ and $r \geq 4\lambda$. In this particular

case, with the notations of (3.4.25), we have :

$$u_{n+1} \leq (1 - \alpha_r \gamma_n) u_n + r^2 \sigma^2 \|Q_r\|^2 \gamma_{n+1}^2,$$

where $u_n = \mathbb{E}\|Z_n\|^2$. The Propositions 3.7.1 (iii) and 3.7.2 (iii) now imply that :

$$\mathbb{E} [\|Z_n\|^2] \leq \mathbb{E} [\|Z_0\|^2] e^{-\alpha_r \Gamma_n} + C_\gamma \frac{2r^2 \|Q_r\|^2}{\alpha_r} \sigma^2 \gamma_n,$$

which, in the end, provide an explicit upper bound of $\mathbb{E}\|Z_n\|^2$ since $Z_n = Q_r^{-1} Z_n$.

A more important issue concerns the rate obtained when $\beta = 1$ and we can remark in the statement of Proposition 3.4.1 that this rate depends on the size of γ and of α_r . In particular, the best rate (of order $\mathcal{O}(n^{-1})$) is obtained when $\gamma \alpha_r > 1$, meaning that α_r must be as large as possible to optimize the performance of the algorithm and we therefore obtain a non-adaptive rate. It is easy to see that $r \mapsto \alpha_r$ increases on $[0, 4\lambda]$ and decreases on $[4\lambda, +\infty)$. It attains its maximal value ($\max_r \alpha_r = 4\lambda$) when $r = 4\lambda$. This maximal value is twice the size of the eigenvalue of the (standard) stochastic gradient descent (SGD). Finally, $\lim_{r \rightarrow +\infty} \alpha_r = 2\lambda$. This limiting value 2λ corresponds to the size of the eigenvalue of the SGD. In other words, the limit $r = +\infty$ in HBF may be seen as an almost identical situation to SGD.

If we compare the rate of convergence of HBF to the one of SGD using the same step size $\gamma_n = \gamma n^{-1}$, we see that choosing a reasonably large r makes it possible to obtain a less stringent condition on γ to recover the (optimal) rate $\mathcal{O}(n^{-1})$. In particular, the rate of the HBF is better when $r \geq 2\lambda$ than the one attained by the SGD. Unfortunately, it seems impossible to obtain an adaptive procedure on the choice of (γ, r) that guarantees the rate $\mathcal{O}(n^{-1})$, unlike the Polyak-Ruppert averaging procedure.

Polynomial memory $r_n = r \Gamma_n^{-1} \rightarrow 0$

This case is more intricate because of the variations with n of the eigenvectors of the matrix C_n defined in (3.4.21).

Proposition 3.4.2. Assume $(\mathbf{H}_{\sigma,2})$. Let $(Z_n)_{n \geq 0}$ be defined by (3.4.20) with $Sp(S) \subset [\lambda, +\infty[$ and $r_n = \frac{r}{\Gamma_n}$.

(i) If $\beta < 1$ and $r > \frac{1+\beta}{2(1-\beta)}$, a constant $c_{\beta,\lambda,r}$ exists such that :

$$\forall n \geq 1 \quad \mathbb{E}\|X_n\|^2 \leq c_{\beta,\lambda,r} \gamma_n,$$

and

$$\forall n \geq 1 \quad \mathbb{E}\|Y_n\|^2 \leq c_{\beta,\lambda} \gamma_n r_n.$$

(ii) If $\beta = 1$, a constant C exists such that :

$$\forall n \geq 1 \quad \mathbb{E}\|X_n\|^2 \leq \frac{C}{\log n}$$

and

$$\forall n \geq 1 \quad \mathbb{E}\|Y_n\|^2 \leq \frac{C}{n \log n}$$

Remark 3.4.2. We can observe that when $\beta < 1$, the rates of the exponential case are preserved under a constraint on r which becomes harder and harder when β is close to 1 : r needs to be

greater than $\frac{1+\beta}{2(1-\beta)}$. Carefully following the proof of this result, we could in fact show that when $1/2 < r < \frac{1+\beta}{2(1-\beta)}$, then $\mathbb{E}\|X_n\|^2 \leq Cn^{-(r-\frac{1}{2})(1-\beta)}$. Since $(r - \frac{1}{2})(1 - \beta) \rightarrow 0$ as $\beta \rightarrow 1$, our upper bound in $(\log n)^{-1}$ related to the case $\beta = 1$ becomes reasonable. Another possible interpretation of the poor convergence rate in that case is that the size of the negative real part of the eigenvalues of C_n is on the order $\frac{1}{n \log n}$, which leads to a contraction of the bias equivalent to $\mathcal{O}\left(e^{-c \sum_1^n \frac{1}{k \log k}}\right)$. Regardless of c , we cannot obtain a polynomial rate of convergence in that case since $\sum_1^n \frac{1}{k \log k} \sim \log \log n$.

Proof of (i) : We study the case $\beta < 1$ here. According to the arguments used in the proof of Proposition 3.4.1 and Subsection 3.4.1, the dynamical system may be reduced to d couples of systems in the form $(x_n^{(i)}, y_n^{(i)})_{n \geq 1}$ so that we only make the proof for a system solution to (3.4.21) under assumption (3.4.22). Another key feature of the polynomial case has been observed in the proof of the a.s. convergence of the algorithm (Theorem 3.3.2) : the study of the rate in the polynomial case involves a normalization of the algorithm with a $\sqrt{r_n}$ -scaling of the Y coordinate. Therefore, we set $\tilde{Z}_n = (\tilde{X}_n, \tilde{Y}_n)$ with $\tilde{X}_n = X_n$ and $\tilde{Y}_n = Y_n/\sqrt{r_n}$. With these notations, we obtain (similar to Lemma 3.3.2) :

$$\tilde{Z}_{n+1} = (I_2 + \tilde{\gamma}_{n+1}\tilde{C}_n)\tilde{Z}_n + \tilde{\gamma}_{n+1}\sqrt{\frac{r_n}{r_{n+1}}}\Sigma_2\Delta N_{n+1}, \quad (3.4.26)$$

with $\tilde{\gamma}_{n+1} = \gamma_{n+1}\sqrt{r_n}$ and :

$$\tilde{C}_n = \begin{pmatrix} 0 & -1 \\ \lambda\sqrt{\frac{r_n}{r_{n+1}}} & \rho_n \end{pmatrix}$$

with

$$\rho_n := \frac{1}{\tilde{\gamma}_{n+1}} \left(\sqrt{\frac{r_n}{r_{n+1}}} - 1 \right) - \frac{r_n}{\sqrt{r_{n+1}}}.$$

Since $r_n = r\Gamma_n^{-1}$, the following expansion holds :

$$\rho_n = \frac{1}{\sqrt{\Gamma_n}} \left(\frac{1}{2\sqrt{r}} - \sqrt{r} \right) + O\left(\frac{\gamma_n}{\Gamma_n^{\frac{3}{2}}}\right). \quad (3.4.27)$$

In particular, for a large enough n , $\rho_n < 0$ if and only if $r > 1/2$. Furthermore, an integer $n_0 \in \mathbb{N}$ exists such that for any $n \geq n_0$, \tilde{C}_n has complex eigenvalues given by :

$$\mu_{\pm}^{(n)} = \frac{1}{2} \left(\rho_n \pm i\sqrt{4\lambda\sqrt{\frac{r_n}{r_{n+1}}} - \rho_n^2} \right) \xrightarrow{n \rightarrow +\infty} \pm i\sqrt{\lambda}.$$

We define the diagonal matrix :

$$\Lambda_n := \begin{pmatrix} \mu_+^{(n)} & 0 \\ 0 & \mu_-^{(n)} \end{pmatrix}$$

and let Q_n be the matrix that satisfies $Q_n^{-1}\Lambda_n Q_n = \tilde{C}_n$. We have :

$$Q_n^{-1} = \begin{pmatrix} 1 & 1 \\ -\mu_+^{(n)} & -\mu_-^{(n)} \end{pmatrix} \quad \text{and} \quad Q_n = \frac{1}{\mu_+^{(n)} - \mu_-^{(n)}} \begin{pmatrix} -\mu_-^{(n)} & -1 \\ \mu_+^{(n)} & 1 \end{pmatrix}.$$

We can now introduce the change of basis brought by Q_n and the new coordinates $Z_n := Q_n \tilde{Z}_n$. We have :

$$\begin{aligned} Z_{n+1} &= Q_{n+1}(I_2 + \tilde{\gamma}_{n+1}\tilde{C}_n)Q_n^{-1}Z_n + \tilde{\gamma}_{n+1}\sqrt{\frac{r_n}{r_{n+1}}}Q_{n+1}\Sigma_2\Delta N_{n+1} \\ &= Q_{n+1}Q_n^{-1}(I_2 + \tilde{\gamma}_{n+1}\Lambda_n)Z_n + \tilde{\gamma}_{n+1}\sqrt{\frac{r_n}{r_{n+1}}}Q_{n+1}\Sigma_2\Delta N_{n+1}. \end{aligned} \quad (3.4.28)$$

We now observe that :

$$Q_{n+1}Q_n^{-1} = I_2 + \Upsilon_n \quad \text{with} \quad \Upsilon_n = (Q_{n+1} - Q_n)Q_n^{-1}$$

and that for n large enough :

$$\|\Upsilon_n\|_\infty \leq C\|Q_{n+1} - Q_n\|_\infty = O(|\mu_+^{(n+1)} - \mu_+^{(n)}|) = O(|\rho_{n+1} - \rho_n| + |\Im(\mu_+^{(n+1)} - \mu_+^{(n)})|).$$

Expansion (3.4.27), the fact that $\sqrt{\frac{r_n}{r_{n+1}}} = 1 + \frac{1}{2}\frac{\gamma_{n+1}}{\Gamma_n} + O\left(\frac{\gamma_{n+1}^2}{\Gamma_n^2}\right)$ and the Lipschitz continuity of $x \mapsto \sqrt{1+x}$ on $[-1/2, +\infty)$ yield :

$$\|\Upsilon_n\|_\infty = O\left(\frac{\gamma_n}{\Gamma_n^{\frac{3}{2}}} + \frac{\gamma_n - \gamma_{n-1}}{\Gamma_n}\right) = O\left(\frac{\gamma_n}{\Gamma_n^{\frac{3}{2}}}\right) = O\left(n^{-\frac{\beta+3}{2}}\right).$$

From the above, we obtain, for any $z \in \mathbb{R}^2$,

$$\|Q_{n+1}Q_n^{-1}(I_2 + \tilde{\gamma}_{n+1}\Lambda_n)z\|^2 \leq \left[\left(1 + \tilde{\gamma}_{n+1}\frac{\rho_n}{2} + O\left(\frac{\gamma_n}{\Gamma_n^{\frac{3}{2}}}\right)\right)^2 + \left(\tilde{\gamma}_{n+1}\Im(\mu_+^{(n)}) + O\left(\frac{\gamma_n}{\Gamma_n^{\frac{3}{2}}}\right)\right)^2 \right] \|z\|^2,$$

which after several computations yields :

$$\|Q_{n+1}Q_n^{-1}(I_2 + \tilde{\gamma}_{n+1}\Lambda_n)z\|^2 \leq \left(1 + \frac{\gamma_{n+1}}{\Gamma_n} \left(\frac{1}{2} - r + o(1)\right)\right) \|z\|^2.$$

Note that a universal constant C (independent of n) exists such that $\|Q_{n+1}\|_\infty \leq C$ and the upper bounds above can be used into (3.4.28) to deduce that :

$$\|Z_{n+1}\|^2 \leq \left(1 + \frac{\gamma_{n+1}}{\Gamma_n} \left(\frac{1}{2} - r\right) + b \left(\frac{\gamma_{n+1}}{\Gamma_n}\right)^2\right) \|Z_n\|^2 + \tilde{\gamma}_{n+1}\Delta M_n + C \frac{\gamma_{n+1}^2}{\Gamma_n} \|\Delta N_{n+1}\|^2, \quad (3.4.29)$$

where $(\Delta M_n)_{n \geq 1}$ is a sequence of martingale increments and b a large enough constant.

When $\gamma_n = \gamma n^{-\beta}$ with $\beta < 1$, the fact that $\Gamma_n = \frac{n^{1-\beta}}{1-\beta} + O(1)$ combined with the upper bound of the variance of the martingale (3.4.22) imply that :

$$\mathbb{E}[\|Z_{n+1}\|^2] \leq \left(1 - \frac{\alpha}{n} + \frac{b}{n^2}\right) \mathbb{E}[\|Z_n\|^2] + C n^{-1-\beta} \quad (3.4.30)$$

where $\alpha := (r - \frac{1}{2})(1 - \beta)$. Under the condition $r > \frac{1+\beta}{2(1-\beta)}$, we observe that :

$$\alpha > \beta.$$

An induction based on Inequality (3.4.30) yields :

$$\begin{aligned} \mathbb{E}[\|Z_{n+1}\|^2] &\leq \mathbb{E}[\|Z_{n_\varepsilon}\|^2] \prod_{\ell=n_\varepsilon}^n \left(1 - \frac{\alpha}{\ell} + \frac{b}{\ell^2}\right) + C \sum_{k=n_\varepsilon+1}^n k^{-1-\beta} \prod_{\ell=k+1}^n \left(1 - \frac{\alpha}{\ell} + \frac{b}{\ell^2}\right) \\ &\leq C n^{-\beta} \end{aligned}$$

where in the second line, we repeated an argument used in the proof of Propositions 3.7.2 and made use of the property $\alpha > \beta$. To conclude the proof, it remains to observe that $\|Q_{n+1}^{-1}\|_\infty \leq C$ regardless of n . \diamond

(ii) When $\beta = 1$, Inequality 3.4.29 leads to :

$$\mathbb{E}[\|Z_{n+1}\|^2] \leq \left(1 - \frac{\alpha}{n \log n} + \frac{b}{n^2 \log n}\right) \mathbb{E}[\|Z_n\|^2] + \frac{C}{n^2 \log n}$$

and a procedure similar to the one used above (given that $\sum_{k=1}^n (k \log k)^{-1} \sim \log(\log n)$) leads to the desired result. $\diamond \square$

3.4.2 The non-quadratic case under exponential memory

The objective of this subsection is to extend the results of the quadratic case to strongly convex functions satisfying $(\mathbf{H}_{SC}(\alpha))$ for a given positive α . As pointed out in Remark 3.2.2, we are not able to obtain neat and somewhat intrinsic results in the polynomial memory case, so we therefore preferred to only consider the exponential memory one.

With the help of Subsection 3.4.1, we can restrain the study to the situation where $d = 1$ and f has a unique minimum in x^* and we denote $\lambda = f''(x^*)$, which is assumed to be positive. We also assume that $\underline{f''} = \inf_{x \in \mathbb{R}} f''(x) > 0$. It is worth noting that in this setting, we are able to obtain some non-asymptotic bounds with some assumptions on λ only. This means that our results do not involve the quantity $\underline{f''}$. To only involve the value of the second derivative in x^* , the main argument is a *power increase* stated in the next lemma.

Lemma 3.4.1. Let $(u_n^{(k)})_{n \geq 0, k \geq 1}$ be a sequence of non-negative numbers satisfying for every integers $n \geq 0$ and $k \geq 1$,

$$u_{n+1}^{(k)} \leq (1 - a_k \gamma_{n+1} + b_k \gamma_{n+1}^2) u_n^{(k)} + C_k (\gamma_{n+1}^2 + \gamma_{n+1} u_n^{(k+1)}) \quad (3.4.31)$$

where $(a_k)_{k \geq 1}$ and $(b_k)_{k \geq 1}$ are sequences of positive numbers. Furthermore, assume that $K \geq 2$ exists and a constant $C > 0$ exists such that :

$$\forall n \geq 1, \quad u_n^{(K)} \leq C \gamma_n. \quad (3.4.32)$$

Then, suppose that $\gamma_n = \gamma n^{-\beta}$ ($\gamma > 0$, $\beta \in (0, 1]$) and that $\underline{a} := \min_{k \leq K} a_k > 0$ and $\bar{b} := \max_{k \leq K} b_k < +\infty$.

(i) If $\beta \in (0, 1)$, a constant $C > 0$ exists such that for every $k \in \{1, \dots, K\}$,

$$\forall n \geq 1, \quad u_n^{(k)} \leq C\gamma_n.$$

(ii) If $\beta = 1$ and $\underline{a}\gamma > 1$, a constant $C > 0$ exists such that for every $k \in \{1, \dots, K\}$,

$$\forall n \geq 2, \quad u_n^{(k)} \leq Cn^{-1}. \quad (3.4.33)$$

Proof Let $K \geq 2$. We proceed by a decreasing induction on $k \in \{1, \dots, K\}$. The initialization is given by (3.4.32). Then, let $k \in \{1, \dots, K-1\}$ and assume that $u_n^{(k+1)} \leq C_{k+1}\gamma_n$ (where C_k is a positive constant that does not depend on n). We can use this upper bound in the second term of the right hand side of (3.4.31) and obtain :

$$u_{n+1}^{(k)} \leq (1 - \underline{a}\gamma_{n+1} + \bar{b}\gamma_{n+1}^2)u_n^{(k)} + C\gamma_{n+1}^2$$

where C is a constant that does not depend on n .

When $\beta < 1$, it follows from Proposition 3.7.1(iii) that :

$$\forall n \geq 1, \quad u_n^{(k)} \lesssim \gamma_n.$$

◇

If $\beta = 1$ and $\underline{a}\gamma > 1$ now, the above control is a consequence of Proposition 3.7.2(iii). This concludes the proof. ◇□

We will apply this lemma to $u_n^{(k)} = \mathbb{E}[|Z_n|^{2k}]$ where Z_n is an appropriate linear transformation of Z_n .

Therefore, we will mainly have to check that Conditions (3.4.31) and (3.4.32) hold.

Proposition 3.4.3. Assume (\mathbf{H}_s) , $(\mathbf{H}_{SC}(\alpha))$ and $(\mathbf{H}_{\sigma,\infty})$ with $p \geq 1$. Let a and b be some positive numbers such that (3.3.8) holds. Then, an integer $K \geq 1$ exists such that for any $p \geq K$:

$$\mathbb{E}[V_n^p(X_n, Y_n)] \leq C_p\gamma_n. \quad (3.4.34)$$

Furthermore, if $r_n = r$ and $\gamma_n = \gamma n^{-\beta}$ with $\beta \in (0, 1)$, then (3.4.34) holds for $p = K = 1$ under $(\mathbf{H}_{\sigma,2})$ instead of $(\mathbf{H}_{\sigma,\infty})$. As a consequence,

$$\mathbb{E}[\|X_n - x^*\|^{2K} + \|Y_n\|^{2K}] \leq C\gamma_n. \quad (3.4.35)$$

Remark 3.4.3. Note that the second assertion (3.4.35) easily follows from Equations (3.3.9) and (3.4.34) and from the fact that under $(\mathbf{H}_{SC}(\alpha))$, a constant c exists such that for all x , $f(x) \geq c\|x\|^2$. Moreover, note that this proposition is not restricted to the exponential memory case. In particular, as suggested in Remark 3.2.2, this Lyapunov approach could lead to some (rough) controls of the quadratic error in the polynomial case when the function is not quadratic.

Proof We begin by the first assertion under Assumption $(\mathbf{H}_{\sigma,\infty})$. Going back to the proof of Lemma 3.3.1 (and to the associated notations), we obtain the existence of some positive a and b such that

$$\begin{aligned} V_{n+1}(X_{n+1}, Y_{n+1}) &\leq V_n(X_n, Y_n) + \gamma_{n+1}\Delta_{n+1} \quad \text{with} \\ \Delta_{n+1} &= -c_{a,b}\|Y_n\|^2 - r_nb\|\nabla f(X_n)\|^2 - br_n\langle \nabla f(X_n), \Delta M_{n+1} \rangle + \Delta R_{n+1} \quad (c_{a,b} > 0). \end{aligned}$$

Denoting the smallest (positive) eigenvalue of $D^2f(x^*)$ by $\underline{\lambda}$, we have :

$$\|\nabla f(x)\|^2 \geq \underline{\lambda}\|x\|^2 \geq C \underline{\lambda}f(x).$$

Following the arguments of the proof of Lemma 3.3.1 once again, we can easily deduce the existence of some positive ε and C such that :

$$\mathbb{E}[\Delta_{n+1}|\mathcal{F}_n] \leq (-\varepsilon + C\gamma_{n+1})r_n V_n(X_n, Y_n) + C\gamma_{n+1}r_n.$$

Using $(\mathbf{H}_{\sigma, \infty})$, we also obtain for every $r \geq 1$:

$$\mathbb{E}[\|\Delta_{n+1}\|^r|\mathcal{F}_n] \leq C_r(1 + V_n^r(X_n, Y_n)).$$

As a consequence, a binomial expansion of $(V_n(X_n, Y_n) + \gamma_{n+1}\Delta_{n+1})^K$ yields :

$$\mathbb{E}[V_{n+1}^K(X_{n+1}, Y_{n+1})|\mathcal{F}_n] \leq (1 - K\varepsilon\gamma_{n+1}r_n + C\gamma_{n+1}^2r_n)V_n^K(X_n, Y_n) + C\gamma_{n+1}^2r_n.$$

Setting $u_n = \mathbb{E}[V_{n+1}^K(X_{n+1}, Y_{n+1})]$, we obtain :

$$u_{n+1} \leq (1 - K\varepsilon\gamma_{n+1}r_n + C\gamma_{n+1}^2r_n)u_n + C\gamma_{n+1}^2r_n.$$

Now, assume that $\gamma_n = \gamma n^{-\beta}$ with $\beta \in (0, 1]$ and successively consider exponential and polynomial cases :

- If $r_n = r$ and $\beta < 1$, the result holds with $K = 1$ by Proposition 3.7.1(iii). \diamond
- If $r_n = r$ and $\beta = 1$, we have to choose K large enough in order that $K\varepsilon\gamma > 1$. In this case, Proposition 3.7.2(iii) yields the result. \diamond
- If $r_n = r/\Gamma_n$ and $\beta < 1$ now, then the above inequality yields the existence of a $\rho > \beta$ and a $n_0 \geq 1$ for K large enough such that :

$$\forall n \geq n_0, \quad u_{n+1} \leq \left(1 - \frac{\rho}{n}\right) u_n + Cn^{-\beta-1}.$$

We have :

$$u_n \leq u_{n_0} \prod_{k=n_0}^n \left(1 - \frac{\rho}{k}\right) + C \sum_{k=n_0+1}^n k^{-\beta-1} \prod_{\ell=k+1}^n \left(1 - \frac{\rho}{\ell}\right).$$

Given that $1 - x \leq \exp(-x)$ and that $\sum_{k=1}^n \frac{1}{k} = \log n + O(1)$, we obtain :

$$u_n \leq Cn^{-\rho} \left(1 + \sum_{k=n_0+1}^n k^{-\beta-1+\rho}\right) \leq Cn^{-\beta}$$

where in the last inequality, we deduced that $-\beta - 1 + \rho > -1$ since $\rho < \beta$.

$\diamond \square$

Proposition 3.4.4. Assume (\mathbf{H}_s) , $(\mathbf{H}_{SC}(\alpha))$ and $(\mathbf{H}_{\sigma, \infty})$ and $r_n = r$ for all $n \geq 1$. Set $\lambda = f''(x^*)$. Then, assume that $\gamma_n = \gamma n^{-\beta}$ with $\beta \in (0, 1]$.

- If $\beta < 1$, then :

$$\mathbb{E}[\|X_n - x^*\|^2] + \mathbb{E}[\|Y_n\|^2] \leq C\gamma_n.$$

— If $\beta = 1$, then for every $\varepsilon > 0$, a constant C_ε exists such that

$$\mathbb{E}[\|X_n - x^*\|^2] \leq C_\varepsilon n^{-(r+\varepsilon-\sqrt{r^2-4\lambda r}1_{r \geq 4\lambda})\gamma \wedge 1}.$$

Proof The starting point is to linearize the gradient :

$$f'(X_n) = \lambda(X_n - x^*) + \phi_n \quad \text{where} \quad \phi_n = (f''(\xi_n) - f''(x^*))(X_n - x^*).$$

Since f'' is Lipschitz continuous, then :

$$|\phi_n| \leq C(X_n - x^*)^2. \quad (3.4.36)$$

Let us begin with the case where the matrix C_n defined in (3.4.21) has real eigenvalues μ_+ and μ_- (given by (3.4.23)). With the notations introduced in (3.4.24),

$$Z_{n+1} = \begin{pmatrix} 1 + \gamma_{n+1}\mu_+ & 0 \\ 0 & 1 + \gamma_{n+1}\mu_- \end{pmatrix} Z_n + r\gamma_{n+1}Q \begin{pmatrix} 0 \\ \phi_n \end{pmatrix} + r\gamma_{n+1}\xi_{n+1}. \quad (3.4.37)$$

As a consequence,

$$\|Z_{n+1}\|^2 \leq (1 + \mu_+\gamma_{n+1})^2\|Z_n\|^2 + C\gamma_{n+1}\|Z_n\|^3 + \gamma_{n+1}^2(\|Z_n\|^4 + \|\Delta N_{n+1}\|^2) + \Delta \mathcal{M}_{n+1}$$

where $(\Delta \mathcal{M}_n)$ is a sequence of martingale increments. Using the elementary inequality $|x| \leq \varepsilon + C_\varepsilon|x|^2$, $x \in \mathbb{R}$ (available for any $\varepsilon > 0$),

$$\|Z_{n+1}\|^2 \leq [(1 + (2\mu_+ + \varepsilon)\gamma_{n+1} + C\gamma_{n+1}^2)]\|Z_n\|^2 + C_\varepsilon\gamma_{n+1}\|Z_n\|^4 + C\gamma_{n+1}^2\|\Delta N_{n+1}\|^2 + \Delta N_{n+1}.$$

Then, by Assumption $(\mathbf{H}_{\sigma,\infty})$ and the fact $\sup_n \mathbb{E}[|Z_n|^r] < +\infty$ for any $r > 1$ (by Proposition 3.4.3 for example), we obtain, for any $k \geq 1$,

$$\mathbb{E}[\|Z_{n+1}\|^{2k}] \leq (1 + k(2\mu_+ + \varepsilon)\gamma_{n+1} + C_k\gamma_{n+1}^2)\mathbb{E}[\|Z_n\|^{2k}] + C_{k,\varepsilon}(\gamma_{n+1}\mathbb{E}[\|Z_n\|^{2k+2}] + \gamma_{n+1}^2).$$

At this stage, we observe that Assumption (3.4.31) is satisfied with $u_n^{(k)} = \mathbb{E}[\|Z_n\|^{2k}]$ and $a_k = k(2\mu_+ + \varepsilon)$. Using Proposition 3.4.3 and Lemma 3.3.1(i), we check that the second assumption of Lemma 3.4.1 also holds. Thus, the result follows in this case from this lemma. \square

3.5 Limit of the rescaled algorithm

In this paragraph, we establish a (functional) Central Limit Theorem when the memory is exponential, *i.e.*, when $r_n = r$ and when $(\mathbf{H}_{SC}(\alpha))$ holds. In particular, f admits a unique minimum x^* . Without loss of generality, we assume that $x^* = 0$.

3.5.1 Rescaling stochastic HBF

We start with an appropriate rescaling by a factor $\sqrt{\gamma_n}$. More precisely, we define a sequence $(\bar{Z}_n)_{n \geq 1}$:

$$\bar{Z}_n = \frac{Z_n}{\sqrt{\gamma_n}} = \begin{pmatrix} \frac{X_n}{\sqrt{\gamma_n}}, \frac{Y_n}{\sqrt{\gamma_n}} \end{pmatrix}.$$

Given that f is \mathcal{C}^2 (and that $x^* = 0$), we “linearize” ∇f around 0 with a Taylor formula and obtain that $\xi_n \in [0, X_n]$ exists such that :

$$\nabla f(X_n) = D^2 f(\xi_n) X_n.$$

Therefore, we can compute that :

$$\bar{Z}_{n+1} = \bar{Z}_n + \gamma_{n+1} b_n(\bar{Z}_n) + \sqrt{\gamma_{n+1}} \begin{pmatrix} 0 \\ \Delta M_{n+1} \end{pmatrix}$$

where b_n is defined by :

$$b_n(z) = \frac{1}{\gamma_{n+1}} \left(\sqrt{\frac{\gamma_n}{\gamma_{n+1}}} - 1 \right) z + \bar{C}_n z, \quad z \in \mathbb{R}^{2d}, \quad (3.5.38)$$

where :

$$\bar{C}_n := \sqrt{\frac{\gamma_n}{\gamma_{n+1}}} \begin{pmatrix} 0 & -I_d \\ r D^2 f(\xi_n) & -r I_d \end{pmatrix}. \quad (3.5.39)$$

It is important to observe that :

$$\frac{1}{\gamma_{n+1}} \left(\sqrt{\frac{\gamma_n}{\gamma_{n+1}}} - 1 \right) = \gamma^{-1}(n+1)^\beta \left[1 + \frac{\beta}{2n} + o(n^{-1}) - 1 \right] = \begin{cases} o(n^{\beta-1}) & \text{if } \beta < 1 \\ \frac{1}{2\gamma} + o(1) & \text{if } \beta = 1 \end{cases} \quad (3.5.40)$$

We associate to the sequence $(\bar{Z}_n)_{n \geq 1}$ a sequence $(\bar{Z}^{(n)})_{n \geq 1}$ of continuous-time processes defined by :

$$\bar{Z}_t^{(n)} = \bar{Z}_n + B_t^{(n)} + M_t^{(n)}, \quad t \geq 0, \quad (3.5.41)$$

where :

$$B_t^{(n)} = \sum_{k=n+1}^{\tilde{N}(n,t)} \gamma_k b_{k-1}(\bar{Z}_{k-1}) + (t - \underline{t}_n) b_{\tilde{N}(n,t)}(\bar{Z}_{\tilde{N}(n,t)}),$$

$$M_t^{(n)} = \sum_{k=n+1}^{\tilde{N}(n,t)} \sqrt{\gamma_k} \begin{pmatrix} 0 \\ \Delta M_k \end{pmatrix} + \sqrt{t - \underline{t}_n} \begin{pmatrix} 0 \\ \Delta M_{\tilde{N}(n,t)+1} \end{pmatrix}.$$

We used the standard notations $\underline{t}_n = \Gamma_{\tilde{N}(n,t)} - \Gamma_n$ above where $N(n, t) = \min \left\{ m \geq n, \sum_{k=n+1}^m \gamma_k > t \right\}$.

To obtain a CLT, we show that $(\bar{Z}^{(n)})_{n \geq 1}$ converges in distribution to a stationary diffusion, following a classical roadmap based on a tightness result and on an identification of the limit as a solution to a martingale problem.

3.5.2 Tightness

The next lemma holds for any sequence of processes that satisfy (3.5.41).

Lemma 3.5.1. Assume that $D^2 f$ is bounded, that $\sup_{k \geq 1} \mathbb{E}[\|\bar{Z}_k\|^2] < +\infty$ and that a $p > 2$ exists such that $\sup_{k \geq 1} \mathbb{E}[\|\Delta M_k\|^p] < +\infty$, then $(\bar{Z}^{(n)})_{n \geq 1}$ is tight for the weak topology induced by the weak convergence on compact intervals.

Proof First, note that $\bar{Z}_0^{(n)} = \bar{Z}_n$, the assumption $\sup_{k \geq 1} \mathbb{E}[\|\bar{Z}_k\|^2] < +\infty$ implies the tightness of $(\bar{Z}_0^{(n)})_{n \geq 1}$ (on \mathbb{R}^{2d}). Then, by a classical criterion (see, *e.g.*, [27, Theorem 8.3]), we deduce that a sufficient condition for the tightness of $(\bar{Z}^{(n)})_{n \geq 1}$ (for the weak topology induced by the uniform convergence on compact intervals) is the following property : for any $T > 0$, for any positive ε and η , a $\delta > 0$ exist and an integer n_0 such that for any $t \in [0, T]$ and $n \geq n_0$,

$$\mathbb{P}\left(\sup_{s \in [t, t+\delta]} \|\bar{Z}_s^{(n)} - \bar{Z}_t^{(n)}\| \geq \varepsilon\right) \leq \eta\delta.$$

We consider $B^{(n)}$ and $M^{(n)}$ separately and begin by the drift term $B^{(n)}$. On the one hand,

$$\mathbb{P}\left(\sup_{s \in [t, t+\delta]} \|B_s^{(n)} - B_t^{(n)}\| \geq \varepsilon\right) \leq \mathbb{P}\left(\sum_{k=N(n,t)}^{N(n,t+\delta)+1} \gamma_k \|b_{k-1}(\bar{Z}_{k-1})\| \geq \varepsilon\right).$$

The Chebyshev inequality and the fact that $\|b_k(z)\| \leq C(1 + \|z\|)$ (where C does not depend on k) yield :

$$\mathbb{P}\left(\sup_{s \in [t, t+\delta]} \|B_s^{(n)} - B_t^{(n)}\| \geq \varepsilon\right) \leq \varepsilon^{-2} \mathbb{E}\left[\left(\sum_{k=N(n,t)}^{N(n,t+\delta)+1} \gamma_k (1 + \|\bar{Z}_{k-1}\|)\right)^2\right]$$

The Jensen inequality and the fact that $\sum_{k=N(n,t)}^{N(n,t+\delta)+1} \gamma_k \leq 2\delta$ when n is large enough imply that a constant C exists such that for large enough n and for a small enough δ :

$$\mathbb{P}\left(\sup_{s \in [t, t+\delta]} \|B_s^{(n)} - B_t^{(n)}\| \geq \varepsilon\right) \leq \varepsilon^{-2} \times C\delta^2(1 + \sup_{k \geq 1} \mathbb{E}[\|\bar{Z}_k\|^2]) \leq \eta\delta$$

◇

We now consider the martingale component $M^{(n)}$: if we denote $\alpha = \sqrt{\frac{t-t_n}{\gamma_{N(n,t)+1}}}$, we have for any $t \geq 0$,

$$M_s^{(n)} = (1 - \alpha)M_{N(n,s)}^{(n)} + \alpha M_{N(n,s)+1}^{(n)}$$

so that $\|M_s^{(n)} - M_t^{(n)}\| \leq \max\{\|M_{N(n,s)}^{(n)} - M_t^{(n)}\|, \|M_{N(n,s)+1}^{(n)} - M_t^{(n)}\|\}$. As a consequence,

$$\mathbb{P}\left(\sup_{s \in [t, t+\delta]} \|M_s^{(n)} - M_t^{(n)}\| \geq \varepsilon\right) \leq \mathbb{P}\left(\sup_{N(n,t)+1 \leq k \leq N(n,t+\delta)+1} \|M_{\Gamma_k}^{(n)} - M_t^{(n)}\| \geq \varepsilon\right)$$

Let $p > 2$ and applying the Doob inequality, the assumption of the lemma leads to :

$$\mathbb{P}\left(\sup_{s \in [t, t+\delta]} \|M_s^{(n)} - M_t^{(n)}\| \geq \varepsilon\right) \leq \varepsilon^{-p} \mathbb{E}\left[\|M_{N(n,t+\delta)+1}^{(n)} - M_t^{(n)}\|^p\right]$$

and the Minkowski inequality yields :

$$\mathbb{P} \left(\sup_{s \in [t, t+\delta]} \|M_s^{(n)} - M_t^{(n)}\| \geq \varepsilon \right) \leq \varepsilon^{-p} \sum_{k=N(n,t)+1}^{N(n,t+\delta)+1} \gamma_k^{\frac{p}{2}} \mathbb{E} [\|\Delta M_k\|^p].$$

Under the assumptions of the lemma, $\mathbb{E}[\|\Delta M_k\|^p] \leq C$. Furthermore, we can use the rough upper bound :

$$\sum_{k=N(n,t)+1}^{N(n,t+\delta)+1} \gamma_k^{\frac{p}{2}} \leq \gamma_n^{\frac{p}{2}-1} \sum_{k=N(n,t)+1}^{N(n,t+\delta)+1} \gamma_k \leq \eta \delta$$

for large enough n . This concludes the proof. $\diamond \square$

Corollary 3.5.1. Let the assumptions of Theorem 3.2.4 hold, then $(\bar{Z}^{(n)})_{n \geq 1}$ is tight.

Proof To prove this result, it is enough to check that the assumptions of Lemma 3.5.1 are satisfied. First, one remarks that the assumptions of Theorem 3.2.4 imply the ones of Theorem 3.2.3(a) so that $\mathbb{E}[\|Z_n - z^*\|^2] \leq C\gamma_n$ (this also holds when $\beta = 1$ since we assume that $\gamma\alpha_r > 1$). As a consequence, $\sup_{k \geq 1} \mathbb{E}[\|\bar{Z}_k\|^2] < +\infty$.

On the other hand, since $(\mathbf{H}_{\sigma, \mathbf{p}})$ holds for a given $p > 2$, we can derive by following the lines of the proof of Proposition 3.4.3 that $\sup_{n \geq 1} \mathbb{E}[V^p(X_n, Y_n)] < +\infty$. As a consequence, $\sup_n \mathbb{E}[f^p(X_n)] < +\infty$ and $(\mathbf{H}_{\sigma, \mathbf{p}})$ leads to :

$$\sup_{n \geq 1} \mathbb{E}[\|\Delta M_n\|^p] \lesssim \sup_n \mathbb{E}[f^p(X_n)] < +\infty.$$

\square

3.5.3 Identification of the limit

Starting from our compactness result above, we now characterize the potential weak limits of $(\bar{Z}^{(n)})_{n \geq 1}$. This step is strongly based on the following lemma.

Lemma 3.5.2. Suppose that the assumptions of Lemma 3.5.1 hold and that :

$$\mathbb{E}[\Delta M_n (\Delta M_n)^t | \mathcal{F}_{n-1}] \xrightarrow{n \rightarrow +\infty} \mathcal{V} \quad \text{in probability,}$$

where σ^2 is a positive symmetric $d \times d$ -matrix. Then, for every C^2 -function $g : \mathbb{R}^{2d} \rightarrow \mathbb{R}$, compactly supported with Lipschitz continuous second derivatives, we have :

$$\mathbb{E}(g(\bar{Z}_{n+1}) - g(\bar{Z}_n) | \mathcal{F}_n) = \gamma_{n+1} \mathcal{L}g(\bar{Z}_n) + R_n^g$$

where $\gamma_{n+1}^{-1} R_n^g \rightarrow 0$ in L^1 and \mathcal{L} is the infinitesimal generator defined in Theorem 3.2.4.

Remark 3.5.1. We recall that \mathcal{L} is the infinitesimal generator of the following stochastic differential equation :

$$d\bar{Z}_t = \bar{H} \bar{Z}_t dt + \Sigma dB_t$$

where : $\bar{H} = \frac{1}{2\gamma} 1_{\{\beta=1\}} I_{2d} + H$ and Σ is defined in Theorem 3.2.4. $(\bar{Z}_t)_{t \geq 0}$ lies in the family of Ornstein-Uhlenbeck processes : on the one hand, the drift and diffusion coefficients being respectively linear and constant, $(\bar{Z}_t)_{t \geq 0}$ is a Gaussian diffusion ; on the other hand, since \bar{H} has negative eigenvalues, $(\bar{Z}_t)_{t \geq 0}$ is ergodic.

Proof C will denote an absolute constant whose value may change from line to line, for the sake of convenience. We use a Taylor expansion between \bar{Z}_n and \bar{Z}_{n+1} and obtain that θ_n exists in $[0, 1]$ such that :

$$\begin{aligned} g(\bar{Z}_{n+1}) - g(\bar{Z}_n) &= \langle \nabla g(\bar{Z}_n), (\bar{Z}_{n+1} - \bar{Z}_n) \rangle + \frac{1}{2} (\bar{Z}_{n+1} - \bar{Z}_n)^T D^2 g(\bar{Z}_n) (\bar{Z}_{n+1} - \bar{Z}_n) \\ &+ \underbrace{\frac{1}{2} (\bar{Z}_{n+1} - \bar{Z}_n)^T (D^2 g(\theta \bar{Z}_n + (1 - \theta) \bar{Z}_{n+1}) - D^2 g(\bar{Z}_n)) (\bar{Z}_{n+1} - \bar{Z}_n)}_{R_{n+1}^{(1)}}. \end{aligned} \quad (3.5.42)$$

We first deal with the remainder term $R_{n+1}^{(1)}$ and observe that (\bar{C}_n) introduced in (3.5.39) is uniformly bounded so that a constant C exists such that $\|b_n(z)\| \leq C\|z\|$. We thus conclude that :

$$\|\bar{Z}_{n+1} - \bar{Z}_n\| \leq C (\gamma_{n+1} \|\bar{Z}_n\| + \sqrt{\gamma_{n+1}} \|\Delta M_{n+1}\|).$$

Using $(\mathbf{H}_{\sigma, p})$, we deduce that for any $\bar{p} \leq p$,

$$\mathbb{E} [\|\bar{Z}_{n+1} - \bar{Z}_n\|^{\bar{p}}] \leq C \gamma_{n+1}^{\frac{\bar{p}}{2}}. \quad (3.5.43)$$

Since $D^2 g$ is Lipschitz continuous and compactly supported, $D^2 g$ is also ε -Hölder for all $\varepsilon \in (0, 1]$. We choose ε such that $2 + \varepsilon \leq p$ and obtain :

$$\mathbb{E} [R_{n+1}] \leq C \mathbb{E} [\|\bar{Z}_{n+1} - \bar{Z}_n\|^{2+\varepsilon}] \leq C \gamma_{n+1}^{1+\frac{\varepsilon}{2}}.$$

We deduce that $\gamma_{n+1}^{-1} R_{n+1}^{(1)} \rightarrow 0$ in L^1 . \diamond

Second, we can express (3.5.40) when $\gamma_n = \gamma n^{-\beta}$ with $\beta \in (0, 1]$ in the following form :

$$\epsilon_n := \frac{1}{\gamma_{n+1}} \left(\sqrt{\frac{\gamma_n}{\gamma_{n+1}}} - 1 \right) - \frac{1}{2\gamma} 1_{\{\beta=1\}} = o(1).$$

Then, given that $D^2 f$ is Lipschitz (and that $x^* = 0$), it follows that :

$$\forall z \in \mathbb{R}^d \times \mathbb{R}^d \quad \left\| b_n(z) - \left(\frac{1}{2\gamma} 1_{\{\beta=1\}} I_{2d} + H \right) z \right\| \leq (\epsilon_n + \|\bar{X}_n\|) \|z\|$$

where $(\epsilon_n)_{n \geq 1}$ is a deterministic sequence such that $\lim_{n \rightarrow +\infty} \epsilon_n = 0$.

Under the conditions of Theorem 3.2.4, we may apply the convergence rates obtained in Theorem 3.2.3 and observe that $\sup_n \mathbb{E}[\|X_n\|^2] \lesssim \gamma_n$, meaning that $\sup_n \mathbb{E}[\|\bar{Z}_n\|^2] < +\infty$. Since $\|\bar{X}_n\| \leq \|\bar{Z}_n\|$, we deduce that :

$$\mathbb{E}[\langle \nabla g(\bar{Z}_n), (\bar{Z}_{n+1} - \bar{Z}_n) \rangle | \mathcal{F}_n] = \gamma_{n+1} \langle \nabla g(\bar{Z}_n), (\frac{1}{4\gamma\sqrt{r}} 1_{\{\beta=1\}} I_{2d} + H) \bar{Z}_n \rangle + R_n^{(2)}$$

where $\gamma_{n+1}^{-1} R_n^{(2)} \rightarrow 0$ in L^1 as $n \rightarrow +\infty$. Let us now consider the second term of the right-hand side of (3.5.42). We have :

$$\mathbb{E}[(\bar{Z}_{n+1} - \bar{Z}_n)^T D^2 g(\bar{Z}_n) (\bar{Z}_{n+1} - \bar{Z}_n) | \mathcal{F}_n] = \gamma_{n+1} \sum_{i,j} D_{y_i y_j}^2 g(\bar{Z}_n) \mathbb{E}[\Delta M_{n+1}^i \Delta M_{n+1}^j | \mathcal{F}_n] + R_n^{(3)}$$

where

$$|\gamma_{n+1}^{-1} R_n^{(3)}| \leq C \gamma_{n+1} \|\bar{Z}_n\|^2 \xrightarrow{n \rightarrow +\infty} 0 \quad \text{in } L^1$$

under the assumptions of the lemma. To conclude the proof, it remains to note that under the assumptions of the lemma for any i and j , $(\mathbb{E}[\Delta M_{n+1}^i \Delta M_{n+1}^j | \mathcal{F}_n])_{n \geq 1}$ is a uniformly integrable sequence that satisfies :

$$\mathbb{E}[\Delta M_{n+1}^i \Delta M_{n+1}^j | \mathcal{F}_n] = \mathcal{V}_{i,j} \quad \text{in probability.}$$

Thus, the convergence also holds in L^1 . The conclusion of the lemma easily follows from the boundedness of $D^2 g$. $\diamond \square$

We are now able to prove Theorem 3.2.4 :

Proof of Theorem 3.2.4, (i) : Note that under the assumptions of Theorem 3.2.4, we can apply Lemma 3.5.1 and Lemma 3.5.2 and obtain that the sequence of processes $(\bar{Z}^{(n)})_{n \geq 1}$ is tight. The rest of the proof is then divided into two steps. In the first one, we prove that every weak limit of $(\bar{Z}^{(n)})_{n \geq 1}$ is a solution of the martingale problem $(\mathcal{L}, \mathcal{C})$ where \mathcal{C} denotes the class of \mathcal{C}^2 -functions with compact support and Lipschitz-continuous second derivatives. Before going further, let us recall that, owing to the Lipschitz continuity of the coefficients, this martingale problem is well-posed, *i.e.*, that existence and uniqueness hold for the weak solution starting from a given initial distribution μ (see, *e.g.*, [47] or [116]).

In a second step, we prove the uniqueness of the invariant distribution related to the operator \mathcal{L} and the convergence in distribution to this invariant measure. We end this proof by showing that $(\bar{Z}^{(n)})$ converges to this invariant distribution, so that the sequence $(\bar{Z}^{(n)})_{n \geq 1}$ converges to a stationary solution of the previously introduced martingale problem. We will characterize this invariant (Gaussian) distribution in the next paragraph.

Step 1 : Let g belong to \mathcal{C} and let $(\mathcal{F}_t^{(n)})_{t \geq 0}$ be the natural filtration of $\bar{Z}^{(n)}$. To prove that any weak limit of $(\bar{Z}^{(n)})_{n \geq 1}$ solves the martingale problem $(\mathcal{L}, \mathcal{C})$, it is enough to show that :

$$\forall t \geq 0, \quad g(\bar{Z}_t^{(n)}) - g(\bar{Z}_0^{(n)}) - \int_0^t \mathcal{L}g(\bar{Z}_s^{(n)}) ds = \mathcal{M}_t^{(n,g)} + \mathcal{R}_t^{(n,g)}$$

where $(\mathcal{M}_t^{(n,g)})_{t \geq 0}$ is an $(\mathcal{F}_t^{(n)})$ -adapted martingale and $\mathcal{R}_t^{(n,g)} \rightarrow 0$ in probability for any $t \geq 0$. We set :

$$\mathcal{M}_t^{(n,g)} = \sum_{k=n+1}^{N(n,t)} g(\bar{Z}_{k+1}) - g(\bar{Z}_k) - \mathbb{E}[g(\bar{Z}_{k+1}) - g(\bar{Z}_k) | \mathcal{F}_{k-1}].$$

By construction, $(\mathcal{M}_t^{(n,g)})_{t \geq 0}$ is an $(\mathcal{F}_t^{(n)})$ -adapted martingale (given that $\mathcal{F}_s^{(n)} = \mathcal{F}_{\underline{s}_n}^{(n)}$) and :

$$\mathcal{R}_t^{(n,g)} = g(\bar{Z}_t^{(n)}) - g(\bar{Z}_{\underline{t}_n}^{(n)}) - \int_{\underline{t}_n}^t \mathcal{L}g(\bar{Z}_s^{(n)}) ds + \int_0^{\underline{t}_n} (\mathcal{L}g(\bar{Z}_{\underline{s}_n}^{(n)}) - \mathcal{L}g(\bar{Z}_s^{(n)})) ds + \sum_{k=n}^{N(n,t)-1} R_k^g$$

where $(R_k^g)_{k \geq 1}$ has been defined in Lemma 3.5.2. Using an argument similar to (3.5.43), we can check that for any $t \geq 0$:

$$\sup_{s \leq t} \mathbb{E}[\|\bar{Z}_s^{(n)} - \bar{Z}_{\underline{s}_n}^{(n)}\|^2] \leq C \sqrt{\gamma_n}.$$

This inequality combined with the Lipschitz continuity of g and its derivatives implies that the first

three terms tend to 0 when $n \rightarrow +\infty$. Now, concerning the last one, the previous lemma yields :

$$\mathbb{E} \left[\left\| \sum_{k=n}^{N(n,t)-1} R_k^g \right\| \right] \leq Ct \sup_{k \geq n} \mathbb{E} [|\gamma_k^{-1} R_k^g|] \xrightarrow{n \rightarrow +\infty} 0.$$

◇

Step 2 : First, let us prove that uniqueness holds for the invariant distribution related to \mathcal{L} . We denote it by $\mu_\infty^{(\beta)}$ below. In this simple setting where the coefficients are linear, we could use the fact that the process, which is solution to the martingale problem, is Gaussian so that any invariant distribution is so. Uniqueness could then be deduced through the characterization of the mean and the variance through the relationship $\int \mathcal{L}f(x) \mu_\infty^{(\beta)}(dx) = 0$ (see next subsection for such an approach). However, at this stage, we prefer to use a more general strategy related to the hypoellipticity of \mathcal{L} (see, *e.g.*, [53] for a similar approach). More precisely, set $L_D := -\langle y, \partial_x \rangle + r\langle D^2 f(x^*)x - y, \partial_y \rangle$ and $\sigma_i := \sum_{j=1}^d \sigma_i^j \partial_{y_j}$, where σ satisfies $\sigma \sigma^t = \mathcal{V}$ (where \mathcal{V} is defined by (3.2.6)). We have assumed that σ is invertible, so that :

$$\text{span}(\sigma_1, \dots, \sigma_d) = \text{span}(\partial_{y_1}, \dots, \partial_{y_d}).$$

Therefore,

$$\text{Lie}(L_D, \sigma_1, \dots, \sigma_d) = \text{Lie}(L_D, \partial_{y_1}, \dots, \partial_{y_d})$$

Now, it is straightforward to check that :

$$\forall i \in \{1, \dots, d\} \quad [L_D, \partial_{y_i}](f) = -\partial_{x_i}(f),$$

and we deduce that $\text{Lie}(L_D, \sigma_1, \dots, \sigma_d) = \text{Lie}(\partial_{x_1}, \dots, \partial_{x_d}, \partial_{y_1}, \dots, \partial_{y_d})$. This means that the Hörmander bracket condition holds at any point z of \mathbb{R}^{2d} , which implies that the process admits a density $(p_t(z, \cdot))_{t \geq 0}$ such that for any $t > 0$, $(z, z') \mapsto p_t(z, z')$, which is smooth on $\mathbb{R}^{2d} \times \mathbb{R}^{2d}$. It is moreover possible to show that these densities are positive, for any $t > 0$, given that the linear vector field is approximately controllable : for any time $T > 0$, any $\eta > 0$ and any couple of initial points (x_0, y_0) and ending points (x_T, y_T) , we can build a function φ such that $\dot{\varphi} \in \mathbb{L}^2$ and such that the controlled trajectory :

$$\begin{cases} \dot{x}(t) &= -y(t) \\ \dot{y}(t) &= r(t)(\nabla U(x(t)) - y(t)) + \sigma \dot{\varphi}, \end{cases} \quad (3.5.44)$$

satisfies : $z_0 = (x_0, y_0)$ and $\|z_T - (x_T, y_T)\| \leq \eta$. This implies the irreducibility of the diffusion and, therefore, the uniqueness of the invariant distribution. We refer to [53] for more details on this controllability problem.

Then, checking that $\mathcal{L}\|x\|^2 \leq \beta - \alpha\|x\|^2$ for positive α and β , it can be classically deduced from the Meyn-Tweedie-type arguments (see [91]) that the process converges locally uniformly, exponentially fast in total variation to $\mu_\infty^{(\beta)}$. For more details, we refer to [88, Theorem 4.4]. Below, we will only use the following corollary : for any bounded Lipschitz-continuous function f , for any compact set K of \mathbb{R}^{2d} ,

$$\sup_{z \in K} |P_t f(z) - \mu_\infty^{(\beta)}(f)| \xrightarrow{t \rightarrow +\infty} 0 \quad (3.5.45)$$

where $(P_t)_{t \geq 0}$ denotes the semi-group related to the (well-posed) martingale problem $(\mathcal{L}, \mathcal{C})$. ◇

Step 3 : Let $(\bar{Z}_{n_k})_{k \geq 1}$ be a (weakly) convergent subsequence of $(\bar{Z}_n)_{n \geq 1}$ to a probability ν . We

have to prove that $\nu = \mu_\infty^{(\beta)}$. To do this, we take advantage of the “shifted” construction of the sequence $(\bar{Z}^{(n)})_{n \in \mathbb{N}}$. More precisely, as a result of construction, for any positive T , a sequence $(\psi(n_k, T))_{k \geq 1}$ exists such that :

$$N(T, \psi(n_k, T)) = n_k.$$

In other words,

$$\bar{Z}_{\underline{T}^{\psi(n_k, T)}}^{(\psi(n_k, T))} = \bar{Z}_{n_k}.$$

At the price of a potential extraction, $(\bar{Z}^{(\psi(n_k, T))})_{k \geq 1}$ is convergent to a continuous process, which is denoted by $Z^{\infty, T}$ below. Given that $\bar{Z}_T^{(n)} - \bar{Z}_{\underline{T}_n}^{(n)}$ tends to 0 as $n \rightarrow +\infty$ in probability, it follows that $Z_T^{\infty, T}$ has distribution ν . However, according to Step 1, $Z^{\infty, T}$ is also a solution to the martingale problem $(\mathcal{L}, \mathcal{C})$ so that for any Lipschitz continuous function f ,

$$\mathbb{E}[f(Z_T^{\infty, T})] - \mu_\infty^{(\beta)}(f) = \int_{\mathbb{R}^{2d}} \left(P_T f(z) - \mu_\infty^{(\beta)}(f) \right) \mathbb{P}_{Z_0^{\infty, T}}(dz).$$

Denote by \mathcal{P} , the set of weak limits of $(\bar{Z}_n)_{n \geq 1}$. \mathcal{P} is tight and as a result of construction, $Z_0^{\infty, T}$ belongs to \mathcal{P} . Thus, for any $\varepsilon > 0$, a compact set K_ε exists such that for any $T > 0$,

$$\left| \int_{K_\varepsilon^c} \left(P_T f(z) - \mu_\infty^{(\beta)}(f) \right) \mathbb{P}_{Z_0^{\infty, T}}(dz) \right| \leq 2\|f\|_\infty \sup_{\mu \in \mathcal{P}} \mu(K_\varepsilon^c) \leq 2\|f\|_\infty \varepsilon.$$

On the other hand,

$$\left| \int_{K_\varepsilon} \left(P_T f(z) - \mu_\infty^{(\beta)}(f) \right) \mathbb{P}_{Z_0^{\infty, T}}(dz) \right| \leq \sup_{z \in K_\varepsilon} |P_T f(z) - \mu_\infty^{(\beta)}(f)|$$

and it follows from Step 2 that the right-hand member tends to 0 as $T \rightarrow +\infty$. From this, we can therefore conclude that for any bounded Lipschitz-continuous function f , a large enough T exists such that :

$$\left| \mathbb{E}[f(Z_T^{\infty, T})] - \mu_\infty^{(\beta)}(f) \right| \leq C_f \varepsilon.$$

Since $\mathbb{E}[f(Z_T^{\infty, T})] = \nu(f)$, it follows that $\nu(f) = \mu_\infty^{(\beta)}(f)$. Finally, the set \mathcal{P} is reduced to a single element $\mathcal{P} = \{\mu_\infty^{(\beta)}\}$, and the whole sequence $(\bar{Z}_n)_{n \geq 1}$ converges to $\mu_\infty^{(\beta)}$.

Before ending this section, let us note that $\mu_\infty^{(\beta)}$ is a Gaussian centered distribution is a simple consequence of Remark 3.5.1. We therefore leave this point to the reader. $\diamond \square$

3.5.4 Limit variance

We end this section on the analysis of the rescaled algorithm with some considerations on the invariant measure $\mu_\infty^{(\beta)}$ involved in Theorem 3.2.4 for the exponential memoried stochastic HBF, *i.e.* when $r_n = r$. As shown in the above paragraph, this invariant measure describes the exact asymptotic variance of the initial algorithm. We now focus on its characterization *i.e.*, on the proof of Theorem 3.2.4(ii). In particular, to ease the presentation, we assume that the covariance matrix \mathcal{V} related to $(\Delta M_{n+1})_{n \geq 1}$ is proportional to the identity matrix :

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\Delta M_{n+1} (\Delta M_{n+1})^t | \mathcal{F}_n \right] = \sigma_0^2 I_d \quad \text{in probability.} \quad (3.5.46)$$

We also assume that $\gamma_n = \gamma n^{-\beta}$ with $\beta < 1$. Then, (i) of Theorem 3.2.4 states that $(\bar{Z}_n)_{n \geq 1}$ weakly converges toward a diffusion process, whose generator \mathcal{L} is the one of an Ornstein-Uhlenbeck process. Assumption (3.5.46) leads to a simpler expression :

$$\mathcal{L}(\phi)(x, y) = -\langle y, \nabla_x \phi \rangle + r \langle D^2(f)(x^*)x - y, \nabla_y \phi \rangle + r^2 \frac{\sigma_0^2}{2} \Delta_y \phi. \quad (3.5.47)$$

A particular feature of Equation (3.5.47) when $\gamma_n = \gamma n^{-\beta}$ is that \mathcal{L} does not depend on β nor γ . The invariant measure $\mu_\infty^{(\beta)}$ is a multivariate Gaussian distribution that may be well described in the basis given by the eigenvectors of the Hessian $D^2(f)(x^*)$. The reduction to d couples of two-dimensional system used in Section 3.4.1 makes it possible to use the spectral decomposition of $D^2(f)(x^*) = P^{-1} \Lambda P$ where P is an orthonormal matrix and Λ a diagonal matrix with positive eigenvalues. The process $(X_n, Y_n) = (P \bar{X}_n, P \bar{Y}_n)$ is therefore centered and Gaussianly distributed asymptotically. This process is associated with $d \times 2$ blockwise independent Ornstein-Uhlenbeck processes, whose generator is now

$$\check{\mathcal{L}}(\phi)(\check{x}, \check{y}) = -\langle \check{y}, \nabla_{\check{x}} \phi \rangle + r \langle \Lambda \check{x} - \check{y}, \nabla_{\check{y}} \phi \rangle + r^2 \frac{\sigma_0^2}{2} \Delta_{\check{y}} \phi,$$

where we used $\text{Tr}(P^t D_y^2 P) = \text{Tr}(D_y^2 P P^t) = \text{Tr}(D_y^2)$ in the last line because $P^t P = I_d$. If we denote $\check{\mu}_\infty^{(\beta)}$ the associated invariant gaussian measure, the tensor structure of $\check{\mathcal{L}}$ leads to

$$\forall i \neq j \quad \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(\beta)}} [\check{x}^{(i)} \check{x}^{(j)}] = \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(\beta)}} [\check{x}^{(i)} \check{y}^{(j)}] = \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(\beta)}} [\check{y}^{(i)} \check{y}^{(j)}] = 0. \quad (3.5.48)$$

Now, using the relationship $\int \check{\mathcal{L}}(\phi) d\check{\mu}_\infty^{(\beta)} = 0$ for some well chosen functions ϕ , we can identify the rest of the covariance matrix. Denote i any integer in $\{1, \dots, d\}$. We chose $\phi(\check{x}, \check{y}) = \frac{\{\check{x}^{(i)}\}^2}{2}$ and obtain that $\check{\mathcal{L}}\left(\frac{\{\check{x}^{(i)}\}^2}{2}\right)(\check{x}, \check{y}) = -\check{x}^{(i)} \check{y}^{(i)}$. It then implies that

$$\mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(\beta)}} [\check{x}^{(i)} \check{y}^{(i)}] = 0. \quad (3.5.49)$$

Picking now $\phi(\check{x}, \check{y}) = \frac{\{\check{y}^{(i)}\}^2}{2}$, we obtain $\check{\mathcal{L}}\left(\frac{\{\check{y}^{(i)}\}^2}{2}\right)(\check{x}, \check{y}) = r \lambda_i \check{x}^{(i)} \check{y}^{(i)} - r \{\check{y}^{(i)}\}^2 + \frac{r^2 \sigma_0^2}{2}$ so that

$$\mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(\beta)}} [\{\check{y}^{(i)}\}^2] = \frac{r \sigma_0^2}{2}. \quad (3.5.50)$$

Finally, we chose $\phi(\check{x}, \check{y}) = \check{x}^{(i)} \check{y}^{(i)}$ and obtain $\check{\mathcal{L}}(\check{x}^{(i)} \check{y}^{(i)})(\check{x}, \check{y}) = -\{\check{y}^{(i)}\}^2 + r \lambda_i \{\check{x}^{(i)}\}^2 - r \check{x}^{(i)} \check{y}^{(i)}$. Therefore, we deduce that :

$$\mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(\beta)}} [\{\check{x}^{(i)}\}^2] = \frac{\sigma_0^2}{2 \lambda_i}. \quad (3.5.51)$$

We can sum-up formulae (3.5.48)-(3.5.51) in $\check{\mu}_\infty^{(\beta)} = \mathcal{N}(0, D_{r, \sigma_0})$ with $D_{r, \sigma_0} = \frac{\sigma_0^2}{2} \begin{pmatrix} \Lambda^{-1} & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & r I_d \end{pmatrix}$.

Since $(\bar{X}_n, \bar{Y}_n) = (P^{-1}X_n, P^{-1}Y_n)$, we deduce that :

$$\mu_\infty^{(\beta)} = \mathcal{N}\left(0, \frac{\sigma_0^2}{2} \begin{pmatrix} \{D^2 f(x^*)\}^{-1} & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & rI_d \end{pmatrix}\right).$$

◇

Theorem 3.2.4- Step size $\gamma_n = \gamma n^{-1}$

This situation is more involved since we can observe that the drift of the limit diffusion is modified according to the size of γ . In particular, the generator \mathcal{L} in that case is shifted from the one above by $\frac{1}{2\gamma}I$ so that :

$$\mathcal{L}(\phi)(x, y) = \frac{1}{2\gamma} [\langle \nabla_x \phi, x \rangle + \langle \nabla_y \phi, y \rangle] - \langle y, \nabla_x \phi \rangle + r \langle D^2 f(x^*)x - y, \nabla_y \phi \rangle + r^2 \frac{\sigma_0^2}{2} \Delta_y \phi.$$

Again, we can use the decomposition $D^2 f(x^*) = P^{-1}\Lambda P$ where P is an orthonormal matrix, and the generator of the rotated process $(X_n, Y_n) = (P\bar{X}_n, P\bar{Y}_n)$ is :

$$\check{\mathcal{A}}(\phi)(x, y) = \left\langle \frac{\check{x}}{2\gamma} - \check{y}, \nabla_{\check{x}} \phi \right\rangle + \left\langle r\Lambda \check{x} + \left(\frac{1}{2\gamma} - r\right) \check{y}, \nabla_{\check{y}} \phi \right\rangle + r^2 \frac{\sigma_0^2}{2} \Delta_{\check{y}} \phi$$

The associated Ornstein-Uhlenbeck process has a unique Gaussian invariant measure $\check{\mu}_\infty^{(1)}$ if and only if $\gamma\alpha_r > 1$ where α_r is the constant defined in the statement of Proposition 3.4.1. The following equations still hold :

$$\forall i \neq j \quad \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(1)}} [\check{x}^{(i)} \check{x}^{(j)}] = \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(1)}} [\check{x}^{(i)} \check{y}^{(j)}] = \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(1)}} [\check{y}^{(i)} \check{y}^{(j)}] = 0. \quad (3.5.52)$$

To determine the rest of the covariance matrix, we follow the same strategy and only address the case $d = 1$ for the sake of convenience. We define : $\sigma_x^2 := \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(1)}} [\check{x}^2]$, $\sigma_y^2 := \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(1)}} [\check{y}^2]$ and $\sigma_{x,y} := \mathbb{E}_{(\check{x}, \check{y}) \sim \check{\mu}_\infty^{(1)}} [\check{x}\check{y}]$.

We start by choosing $\phi(\check{x}, \check{y}) = \frac{\check{x}^2}{2}$ and obtain $\check{\mathcal{A}}(\phi)(x, y) = \frac{\check{x}^2}{2\gamma} - \check{x}\check{y}$. Therefore, we deduce that :

$$2\gamma \sigma_{x,y} = \sigma_x^2 \quad (3.5.53)$$

Now we pick $\phi(\check{x}, \check{y}) = \frac{\check{y}^2}{2}$ and obtain $\check{\mathcal{A}}(\phi)(x, y) = r\lambda \check{x}\check{y} + \left(\frac{1}{2\gamma} - r\right) \check{y}^2 + \frac{r^2 \sigma_0^2}{2}$ so that :

$$\left(r - \frac{1}{2\gamma}\right) \sigma_y^2 = r\lambda \sigma_{x,y} + \frac{r^2 \sigma_0^2}{2}. \quad (3.5.54)$$

Finally, the function $\phi(\check{x}, \check{y}) = \check{x}\check{y}$ yields $\check{\mathcal{A}}(\phi)(x, y) = \check{x}\check{y} \left(\frac{1}{\gamma} - r\right) - \check{y}^2 + r\lambda \check{x}^2$, which implies :

$$\sigma_y^2 = r\lambda \sigma_x^2 + \left(\frac{1}{\gamma} - r\right) \sigma_{x,y} \quad (3.5.55)$$

We are led to the introduction of :

$$\check{\alpha}_- = 1 - \sqrt{1 - \frac{4\lambda}{r}} \quad \text{and} \quad \check{\alpha}_+ = 1 + \sqrt{1 - \frac{4\lambda}{r}},$$

which leads to :

$$\sigma_x^2 = \sigma_0^2 \frac{2\lambda r \gamma^3}{(\gamma r - 1)(2\lambda\gamma - \check{\alpha}_-)(2\lambda\gamma - \check{\alpha}_+)}, \quad \sigma_y^2 = \sigma_0^2 \frac{\lambda r \gamma (2\lambda r \gamma^2 - r\gamma + 1)}{(\gamma r - 1)(2\lambda\gamma - \check{\alpha}_-)(2\lambda\gamma - \check{\alpha}_+)},$$

and

$$\sigma_{x,y} = \sigma_0^2 \frac{\lambda r \gamma^2}{(\gamma r - 1)(2\lambda\gamma - \check{\alpha}_-)(2\lambda\gamma - \check{\alpha}_+)}.$$

◇□

3.6 Numerical experiments

In this short paragraph, we briefly investigate the behavior of several algorithms, widely used in the field of stochastic approximation. In particular, we are interested in the convergence rates of each algorithm, as well as their behavior in the case of non-convex potential f with multiple wells, to illustrate both Theorem 3.2.3 and Theorem 3.2.2.

Convergence rates We are first concerned by the typical behavior of the heavy ball stochastic approximation algorithm in the convex case. In particular, we are interested in the role played by the parameter r that varies, both in the polynomial case and in the exponential case. Figure 3.1 represents the logarithmic loss of the algorithms with respect to the logarithm of the number of iterations in the 1 dimensional case with $f(x) = \frac{x^2}{2}$. The step size used is $\gamma_k = k^{-1}$. We immediately

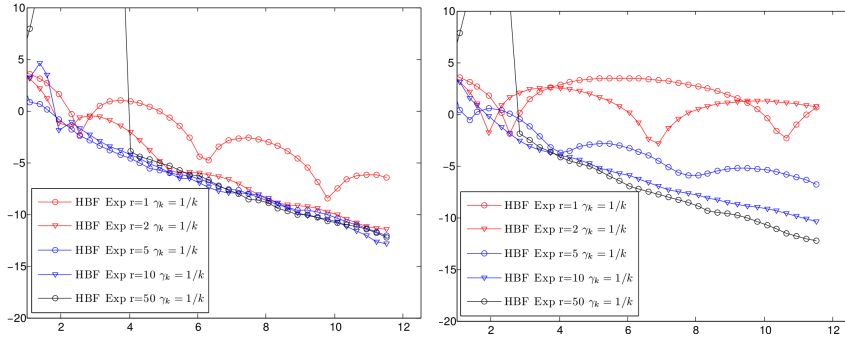


FIGURE 3.1 – Evolution of $\log(f(X_k))$ with respect to $\log(k)$. Left : Exponential memory. Right : Polynomial memory.

observe that for small values of r (that correspond to a long-term memory case), the algorithm possesses a lengthy oscillating behavior, which is a feature of second-order algorithms with a very mild damping effect. This phenomenon has also been observed in previous works (see, *e.g.*, [11] and the references therein). We also observe that the use of an excessively large value of r (say, when r is greater than 10) creates a numerical instability at the beginning of the iterations. This could be fixed by using a supplementary truncating trick introduced in [83]. Finally, the obtained rates are better (from a numerical point of view) when r is chosen at around 5 in the exponential case, and at around 10 in the polynomial case, mainly because of the oscillations that deteriorate the convergence when r is too small.

Figure 3.2 then compares several stochastic optimization algorithms in the toy example $f(x) = |x|^p/p$: the standard Robbins-Monro stochastic gradient descent introduced in [112] (SGD) and several second order algorithms: the “optimal” Ruppert-Polyak averaging algorithm (see [110, 113]), the Nesterov accelerated gradient descent [96] adapted in the stochastic framework in a straightforward way using an unbiased evaluation of the gradient in each iteration, and the recent SAGE method introduced in [67]. Note that the Ruppert-Polyak averaging algorithm is used according to the recommendation of [10] with $\gamma_k = k^{-1/2}$.

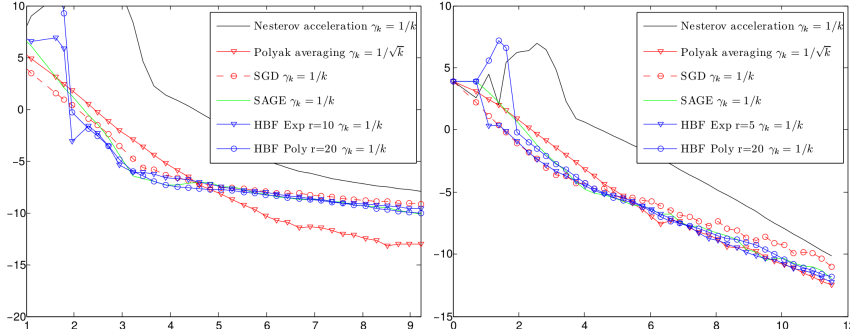


FIGURE 3.2 – Evolution of $\log(f(X_k))$ with respect to $\log(k)$ with $f(x) = |x|^p/p$. Left : Convex case $p = 4$. Right : Strongly convex case $p = 2$.

The first elementary remark is that the rate is of course deteriorated by the loss of strong convexity (left side, Figure 3.2). In this case, the Ruppert-Polyak averaging outperforms other methods and attains the $O(1/\sqrt{n})$ minimax rate (see [95]). When f is strongly convex, the second-order algorithms then all share an equivalent efficiency with, apparently a $O(1/n)$ convergence rate. This corresponds to (ii) of Theorem 3.2.3 when the Hessian at the critical point is sufficiently large to make this minimax optimal rate possible. Nevertheless, the ability of the stochastic heavy ball in a more general situation may deserve further numerical investigation, which is beyond the scope of this paper. The SGD seems to be a little bit less effective in the strongly convex case. Finally, the Nesterov adaptation to the stochastic case does not lead to an efficient algorithm (in comparison to the other methods tested). However, this remark should be balanced by the fact that we did not use the Lan adaptation of the Nesterov accelerated gradient descent introduced in [80]. It appears that this modification that consists in an addition of an intermediary point in the NAGD seems important to optimize the behavior of the algorithm in the stochastic case.

Non-convex case In this paragraph, we investigate the ability of the stochastic algorithm to avoid local traps and, in particular, we focus on the behavior of second order algorithms that may be an intermediary step towards global optimization methods such as simulated annealing. For this purpose, we defined f as :

$$\forall x \in \mathbb{R} \quad f(x) = ax^4 + b(x-1)^2.$$

with $a = 1/40$ and $b = -1/5$. These values have been fixed to guarantee the numerical stability of the stochastic procedures, but the results we obtained may be replicated for other values. The values of a and b above yield a double-well potential with a global minimizer of f of around

$x^* \simeq -4.9$, although f has a local trap on the positive part at around $x_+ \simeq 4$. The function f is represented on the top left of Figure 3.3.

We used $\gamma_k = k^{-1}$ for all of the methods and we varied the initialization point of each algorithm from -10 to 10 with 100 Monte-Carlo replications. For each simulation, we arbitrarily stopped the evolution of the algorithm after $T = 10^4$ iterations, and considered that optimization was successful when $|x_T - x^*| \leq 1$. This criterion may be replaced by a more stringent inequality, at the price of an increase of T , without really changing the main conclusions below.

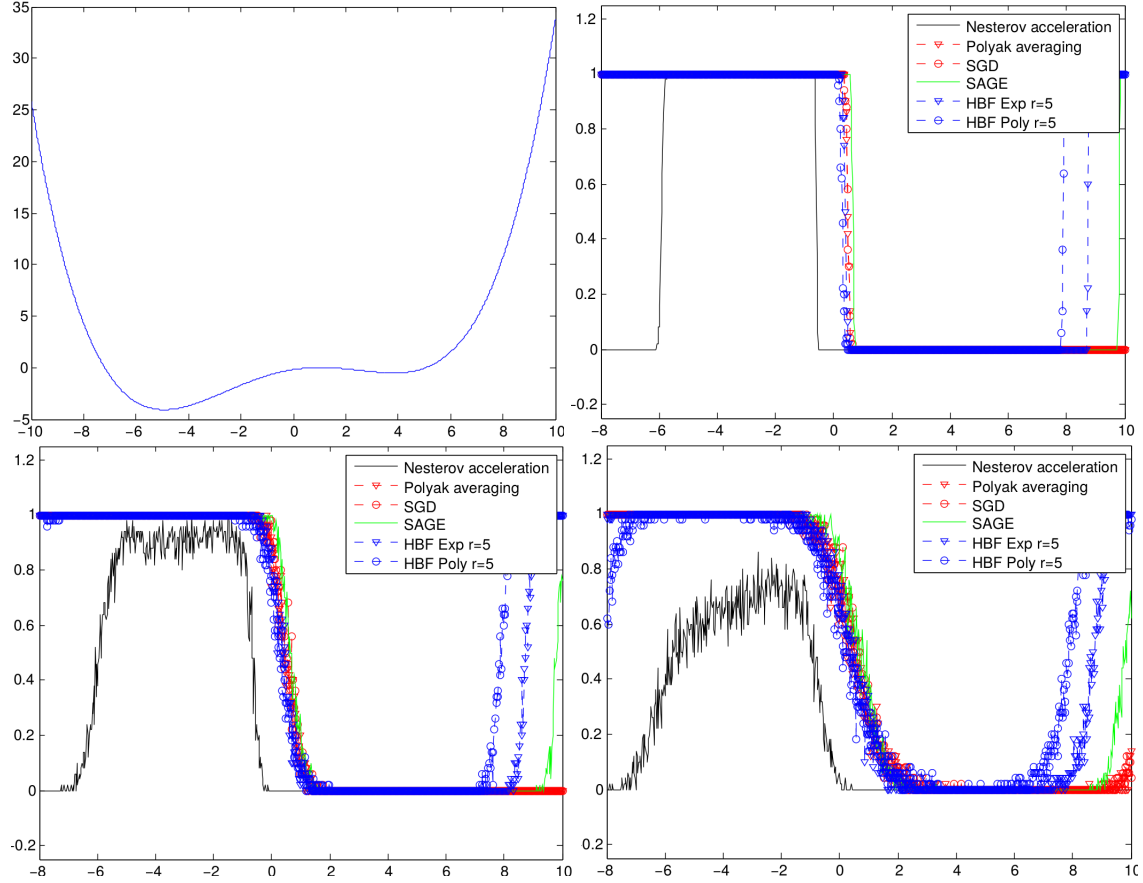


FIGURE 3.3 – Top left : function f to be minimized. Top right : probability of success of the stochastic algorithms with respect to the initialization point with small variance : $\sigma = 0.1$. Bottom left : $\sigma = 1$. Bottom right : $\sigma = 2$.

Performances are reported in Figure 3.3. We observe that both SGD and Ruppert-Polyak algorithms have the same behavior. This fact is absolutely clear because Polyak averaging is built with a Cesaro average of SGD. The target convergence point of SGD and of Polyak averaging are thus the same. We can also note that in the almost no noise setting, the basin of attraction of x^* for SGD may be roughly approximated by $]-\infty, 1]$. Nevertheless, both SAGE and HBF seem to behave better behaviour with a somewhat larger basin of attraction : in particular, it is possible

σ	SGD	AV SGD	SAGE	NAGD	HBF Poly r=5	HBF Expo r=5
0.1	0.47	0.47	0.49	0.29	0.58	0.52
1	0.47	0.47	0.49	0.27	0.58	0.55
2	0.47	0.47	0.49	0.20	0.58	0.54

TABLE 3.1 – Average rate of success of each stochastic algorithm with a uniformly sampled initialization over $[-10; 10]$ when σ varies.

Exp 1	Exp 2	Exp 5	Exp 10	Poly 1	Poly 2	Poly 5	Poly 10
0.51	0.53	0.55	0.58	0.26	0.43	0.58	0.50

TABLE 3.2 – Average rate of success of heavy ball stochastic algorithm for several values of r , when $\sigma = 1$ and the initialization point is sampled uniformly over $[-10; 10]$.

to start from an initialization point $x_1 = 8$ and still obtain convergence of SAGE or HBF towards x^* . This last point is clearly impossible with SGD. The same conclusions hold for different values of σ (see Figure 3.3, bottom left and right). Finally, we observe that NAGD does not present very good behavior : the probability of failure when the algorithm is initialized at -4 is lower than 1 for $\sigma = 1$ or $\sigma = 2$.

We can calculate a more quantitative indicator of this behavior with the computation of the average rate of success of each algorithm when the initialization point is sampled uniformly over $[-10; 10]$. Table 3.1 seems to indicate that the stochastic heavy ball leads to a better exploration of the state space, in particular, with reasonable values of r (see Table 3.2). These conclusions should be understood as numerical observations of experimental results on this particular type of synthetic case, but we do not have any theoretical arguments to strengthen these final observations at this time.

3.7 Technical results

3.7.1 Standard tools of stochastic algorithms

We recall below a standard version of the so-called Robbins-Siegmund Theorem (see *e.g.* [44]) :

Theorem 3.7.1. Given a filtration \mathcal{F}_n and four positive, integrable and \mathcal{F}_n -adapted sequences $(\alpha_n)_n, (\beta_n)_n, (U_n)_n$ and $(V_n)_n$ satisfying :

— (i) $(\alpha_n)_n, (\beta_n)_n, (U_n)_n$ are predictable sequences.

— (ii) $\sup_{\omega} \prod_n (1 + \alpha_n(\omega)) < \infty, \sum_n \mathbb{E}(\beta_n) < \infty$.

— (iii) $\forall n \in \mathbb{N}$,

$$\mathbb{E}(V_{n+1} | \mathcal{F}_n) \leq V_n(1 + \alpha_{n+1}) + \beta_{n+1} - U_{n+1}$$

Then :

(i) V_n converges to V_{∞} in L^1 and $\sup_n \mathbb{E}[V_n] < \infty$.

(ii) $\sum_n \mathbb{E}(U_n) < \infty, \sum_n U_n < \infty$ a.s.

Step sizes $\gamma_n = \gamma n^{-\beta}$ **with** $\beta < 1$

Proposition 3.7.1. For any positive values $a > 0$ and $b > 0$, for any $\beta \in (0, 1)$ and any sequence $(\gamma_n)_{n \geq 1}$ defined by $\gamma_n = \gamma n^{-\beta}$, one has :

- (i) - a If $\beta < 1/2$, then $\sum_{k=1}^n a\gamma_k - b\gamma_k^2 \geq \frac{a\gamma}{1-\beta} n^{1-\beta} - \frac{b\gamma^2}{1-2\beta} n^{1-2\beta}$
- (i) - b If $\beta > 1/2$, then $\sum_{k=1}^n a\gamma_k - b\gamma_k^2 \geq \frac{a\gamma}{1-\beta} n^{1-\beta} - \frac{b\gamma^2}{2\beta-1}$
- (i) - c If $\beta = 1/2$, then $\sum_{k=1}^n a\gamma_k - b\gamma_k^2 \geq \frac{a\gamma}{1-\beta} n^{1-\beta} - b\gamma^2 \log n$
- (ii) An integer n_0 exists such that $\forall n \geq n_0 \quad \sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l)^2 \leq \frac{2}{a} \gamma_{n+1}$
- (iii) An integer n_0 exists such that $\forall n \geq n_0 \quad \sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) \leq \frac{2}{a} \gamma_{n+1}$

Proof : The upper bounds involved in (i) - a, (i) - b and (i) - c are straightforward. \diamond

Proof of (ii) : Using Γ_n introduced in the beginning of Section 3.2, we write :

$$\begin{aligned} \sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l)^2 &\leq \sum_{k=1}^n \gamma_k^2 e^{-a \sum_{k+1}^n \gamma_l} \\ &= \sum_{k=1}^n \gamma_k^2 e^{-a\Gamma_n + a\Gamma_k} \leq \gamma^2 e^{-a\Gamma_n} \sum_{k=1}^n k^{-2\beta} e^{\frac{a\gamma}{1-\beta} k^{1-\beta}} \end{aligned}$$

The function $x \mapsto x^{-2\beta} e^{\frac{a\gamma}{1-\beta} x^{1-\beta}}$ being increasing for $x \geq c_{a,\gamma,\beta}$, we then obtain, considering an integer $t > c_{a,\gamma,\beta}$:

$$\sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l)^2 \leq \gamma^2 e^{-a\Gamma_n} \left(C_t + \int_t^n x^{-2\beta} e^{\frac{a\gamma}{1-\beta} x^{1-\beta}} dx \right).$$

We can write $x^{-2\beta} e^{Kx^{1-\beta}} = \left(e^{Kx^{1-\beta}} \right)' x^{-\beta} K^{-1} (1-\beta)^{-1}$ and integrating by parts, we obtain for a large enough n :

$$\sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l)^2 \leq \gamma^2 e^{-a\Gamma_n} \left(C_t + \frac{e^{a\Gamma_n}}{a\gamma} n^{-\beta} \right) \leq \frac{2}{a} \gamma_n.$$

\diamond

Proof of (iii) : We only deal with $\beta < 1/2$, which is the most involved situation. Using Γ_n and $\Gamma_n^{(2)}$

introduced in the beginning of Section 3.2, we write :

$$\begin{aligned}
\sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) &\leq \sum_{k=1}^n \gamma_k^2 e^{-a\Gamma_n + a\Gamma_k + b\Gamma_n^{(2)} - b\Gamma_k^{(2)}} \\
&\leq e^{-a\Gamma_n + b\Gamma_n^{(2)}} \sum_{k=1}^n \gamma_k^2 e^{a\Gamma_k - b\Gamma_k^{(2)}} \\
&\leq \gamma^2 e^{-a\Gamma_n + b\Gamma_n^{(2)}} \sum_{k=1}^n k^{-2\beta} e^{\frac{a\gamma}{1-\beta} k^{1-\beta} - \frac{b\gamma^2}{1-2\beta} k^{1-2\beta}}
\end{aligned}$$

The function $x \mapsto x^{-2\beta} e^{\frac{a\gamma}{1-\beta} x^{1-\beta} - \frac{b\gamma^2}{1-2\beta} x^{1-2\beta}}$ being increasing for $x \geq c_{a,b,\gamma,\beta}$, we then obtain considering an integer $t > c_{a,b,\gamma,\beta}$:

$$\begin{aligned}
\sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) &\leq \gamma^2 e^{-a\Gamma_n + b\Gamma_n^{(2)}} \left(\sum_{k=1}^t k^{-2\beta} e^{\frac{a\gamma}{1-\beta} k^{1-\beta} - \frac{b\gamma^2}{1-2\beta} k^{1-2\beta}} + \int_t^n x^{-2\beta} e^{\frac{a\gamma}{1-\beta} x^{1-\beta} - \frac{b\gamma^2}{1-2\beta} x^{1-2\beta}} dx \right) \\
&\leq \gamma^2 e^{-a\Gamma_n + b\Gamma_n^{(2)}} \left(C_t + \int_t^n x^{-2\beta} e^{\frac{a\gamma}{1-\beta} x^{1-\beta} - \frac{b\gamma^2}{1-2\beta} x^{1-2\beta}} dx \right) \\
&\leq \gamma^2 e^{-a\Gamma_n + b\Gamma_n^{(2)}} (C_t \\
&\quad + \int_t^n x^{-\beta} \left[\frac{3}{2} \frac{a\gamma x^{-\beta} - b\gamma^2 x^{-2\beta}}{a\gamma} + \frac{3b\gamma x^{-2\beta} - a\gamma x^{-\beta}}{2} \right] e^{\frac{a\gamma}{1-\beta} x^{1-\beta} - \frac{b\gamma^2}{1-2\beta} x^{1-2\beta}} dx)
\end{aligned}$$

Now choosing $t \geq (3b/a)^{\beta-1}$ yields $3b\gamma x^{-2\beta} \leq a\gamma x^{-\beta}$ for any $x \geq t$. Integrating by parts, we obtain :

$$\begin{aligned}
\sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) &\leq \gamma^2 e^{-a\Gamma_n + b\Gamma_n^{(2)}} \left(C_t + \frac{n^{-\beta}}{a\gamma} e^{-a\Gamma_n + n\Gamma_n^{(2)}} \right) \\
&\leq \frac{\gamma n^{-\beta}}{a} + \gamma^2 C_t e^{-a\Gamma_n + b\Gamma_n^{(2)}}.
\end{aligned}$$

Then, choosing n_0 large enough (that depends on a, b, γ and β), we deduce that :

$$\forall n \geq n_0 \quad \sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) \leq \frac{2}{a} \gamma_n.$$

◇
□

Step sizes $\gamma_n = \gamma n^{-1}$

Proposition 3.7.2. For any positive values $a > 0$ and $b > 0$ and any sequence $(\gamma_n)_{n \geq 1}$ defined by $\gamma_n = \gamma n^{-1}$, we have :

$$(i) \quad \sum_{k=1}^n a\gamma_k - b\gamma_k^2 \geq a \log n - b\pi^2/6$$

$$\begin{aligned}
(ii) \quad & \sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l)^2 \leq C_\gamma \begin{cases} \frac{1}{a\gamma-1} n^{-1} & \text{if } a\gamma > 1 \\ \log n n^{-1} & \text{if } a\gamma = 1 \\ \frac{1}{1-a\gamma} n^{-a\gamma} & \text{if } a\gamma < 1 \end{cases} \\
(iii) \quad & \sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) \leq C_{\gamma,b} \begin{cases} \frac{1}{a\gamma-1} n^{-1} & \text{if } a\gamma > 1 \\ \log n n^{-1} & \text{if } a\gamma = 1 \\ \frac{1}{1-a\gamma} n^{-a\gamma} & \text{if } a\gamma < 1 \end{cases} \\
(iv) \quad & \text{For any } \epsilon > 0, a > 0 \text{ and } b > 0 : \sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^{1+\epsilon}) \leq \frac{2e^{b\Gamma_\infty^{(1+\epsilon)}}}{a}.
\end{aligned}$$

Proof : The upper bounds involved in (i) and (ii) are straightforward. \diamond

Proof of (iii) : The situation is easier than the one involved in point (ii) of Proposition 3.7.1 because in that case, we have :

$$\forall n \geq 1 \quad \Gamma_n^{(2)} \leq \gamma^2 \pi^2 / 6.$$

Therefore, we can repeat the computations above and get :

$$\begin{aligned}
\sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) & \leq \sum_{k=1}^n \gamma_k^2 e^{-a\Gamma_n + a\Gamma_k + b\Gamma_n^{(2)} - b\Gamma_k^{(2)}} \\
& \leq e^{-a\Gamma_n + b\gamma^2 \pi^2 / 6} \sum_{k=1}^n \gamma_k^2 e^{a\Gamma_k} \\
& \leq \gamma^2 e^{b\gamma^2 \pi^2 / 6} n^{-a\gamma} \sum_{k=1}^n k^{-2+a\gamma}.
\end{aligned}$$

We then deduce that :

$$\sum_{k=1}^n \gamma_k^2 \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^2) = \gamma^2 e^{b\gamma^2 \pi^2 / 6} \begin{cases} \frac{1}{a\gamma-1} n^{-1} & \text{if } a\gamma > 1 \\ \log n n^{-1} & \text{if } a\gamma = 1 \\ \frac{1}{1-a\gamma} n^{-a\gamma} & \text{if } a\gamma < 1 \end{cases}$$

\diamond

Proof of (iii) : We follow the same guideline : remark that $(\Gamma_n^{(1+\epsilon)})_{n \geq 1}$ is a bounded sequence and write

$$\begin{aligned}
\sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l + b\gamma_l^{1+\epsilon}) & \leq \sum_{k=1}^n \frac{\gamma}{k+1} e^{-a\gamma \log n + a\gamma \log k + b\Gamma_n^{(1+\epsilon)}} \\
& \leq \gamma e^{b\Gamma_\infty^{(1+\epsilon)}} n^{-a\gamma} \int_1^n x^{a\gamma-1} dx \\
& \leq \frac{e^{b\Gamma_\infty^{(1+\epsilon)}}}{a}.
\end{aligned}$$

\square

3.7.2 Expectation of the supremum of the square of sub-Gaussian random variables

We consider a sequence of independent random variables $(\xi_i)_{i \geq n}$ of \mathbb{R}^d such that each coordinate satisfies a sub-Gaussian assumption $(\mathbf{H}_{\text{Gauss}, \sigma})$:

$$\forall \lambda \in \mathbb{R} \quad \forall j \in \{1, \dots, d\} \quad \forall i \geq n \quad \log \mathbb{E} \left[e^{\lambda \xi_i^j} \right] \leq \lambda^2 \frac{\sigma^2}{2}, \quad (3.7.56)$$

where σ^2 is a variance factor. If $(\gamma_k)_{k \geq n}$ is a decreasing sequence in $\ell^2(\mathbb{N})$, we are looking for an upper bound of :

$$m_n^* = \mathbb{E} \left[\sup_{k \geq n} \left\{ \gamma_k^2 \|\xi_k\|^2 \right\} \right]. \quad (3.7.57)$$

For any $\nu > 0$ and any decreasing sequence $\gamma_n \sim \gamma n^{-\nu}$, we establish the following result (useful for Theorem 3.3.2).

Theorem 3.7.2. If each coordinate ξ_i^j is absolutely continuous w.r.t. the Lebesgue measure and satisfies $(\mathbf{H}_{\text{Gauss}, \sigma})$, then :

$$m_n^* \lesssim \sigma^2 d \gamma_n^2 \log(\gamma_n^{-2}),$$

where \lesssim refers to an inequality up to a universal constant.

We begin with a preliminary lemma.

Lemma 3.7.1. Assume that X is a real random variable that satisfies $(\mathbf{H}_{\text{Gauss}, \sigma})$ with median 0 :

$$\mathbb{P}(X > 0) = \mathbb{P}(X < 0) = \frac{1}{2}.$$

Then, we can find $Y \sim \mathcal{N}(0, \sigma^2)$ on the same probability space and c large enough s.t.

$$|X| \leq c|Y| \quad a.s.$$

Proof :

We use a coupling argument. We denote F_X as the cumulative distribution function :

$$F_X(t) = \int_{-\infty}^t f_X(u) du = \mathbb{P}[X \leq t].$$

Similarly, we also denote Ψ_{σ^2} as the cumulative distribution function of a Gaussian random variable $\mathcal{N}(0, \sigma^2)$:

$$\Psi_{\sigma^2}(t) = \int_{-\infty}^t \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi}\sigma} dx = \mathbb{P}[\mathcal{N}(0, \sigma^2) \leq t].$$

Our assumption on the distribution on X shows that the generalized inverse of F_X (denoted F_X^{-1}) exists and if \mathcal{U} is a uniform random variable between on $[0, 1]$, then $X \sim F_X^{-1}(\mathcal{U})$. We now consider the random variable $Y \sim F_{\sigma^2}^{-1}(\mathcal{U})$ built with the same realization of \mathcal{U} . Of course, Y is distributed according to a Gaussian random variable $\mathcal{N}(0, \sigma^2)$.

We need to show that a sufficiently large $c > 0$ exists such that $|X| \leq c|Y|$, that is :

$$|F_X^{-1}(u)| \leq c |\Psi_{\sigma^2}^{-1}(u)|. \quad (3.7.58)$$

Using the fact that F_X is an increasing function, and letting $u = \Psi_{\sigma^2}(y)$, it is then equivalent to

show that :

$$\forall y \in \mathbb{R} \quad F_X(-c|y|) \leq \Psi_{\sigma^2}(|y|) \leq F_X(c|y|) \quad (3.7.59)$$

We now study two different situations for y . If $y = 0$, then Inequality (3.7.59) holds since the median of X is 0. If $|y| \leq \eta$ is close to 0, the same inequality is satisfied with a first-order Taylor expansion. For example, the right hand side reads :

$$F_X(c|y|) \sim \frac{1}{2} + \int_0^{c|y|} f_X(u) du \geq \frac{1}{2} + cf_X(0)|y| + o(|y|),$$

which is greater than $\Psi_{\sigma^2}(|y|)$ for c large enough. Hence, we deduce that Inequality (3.7.59) holds around 0.

Now, we assume that $|y| > \eta > 0$, the desired upper bound (3.7.59) is equivalent to :

$$1 - F_X(c|y|) \leq 1 - \Psi_{\sigma^2}(|y|).$$

The Chernoff bound associated with the sub-Gaussian assumption ($\mathbf{H}_{\mathbf{Gauss}, \sigma}$) on the distribution of X implies that :

$$\mathbb{P}(X > c|y|) \leq e^{\inf_{\lambda > 0} \{\lambda^2 \sigma^2 / 2 - \lambda c|y|\}} = e^{-\frac{c^2 |y|^2}{2\sigma^2}}.$$

At the same time, the lower bound of the Gaussian tail is given by :

$$1 - \Psi_{\sigma^2}(c|y|) \geq \frac{e^{-|y|^2/2\sigma^2}}{\sqrt{2\pi}\sigma} [|y|^{-1} - |y|^{-3}] \geq \kappa(\delta) e^{-|y|^2/2\sigma^2},$$

with $\kappa(\delta)$ a constant independent of $|y| \geq \delta$. Hence, the right hand side of (3.7.59) holds for a large enough c (independent on σ^2). A symmetry argument permits to conclude for the left hand side of (3.7.59).

Inequality (3.7.59) being equivalent to (3.7.58), the conclusion of the proof follows. \square

We are now looking at to the proof of Theorem 3.7.2.

Proof of Theorem 3.7.2 :

We will shift all of the coordinates of the random variables $(\xi_i)_{i \geq n}$ by their corresponding medians. Assuming ($\mathbf{H}_{\mathbf{Gauss}, \sigma}$), the coordinates $(\xi_i^j)_{1 \leq j \leq d}$ are centered and have a second-order moment upper bounded by σ^2 (see [115], for example) :

$$\forall i \geq n \quad \forall j \in \{1, \dots, d\} \quad \mathbb{E}[\{\xi_i^j\}^2] \leq \sigma^2.$$

The Tchebychev inequality implies that each median m_i^j of the random variables ξ_i^j are bounded by :

$$\forall i \geq n \quad \forall j \in \{1, \dots, d\} \quad |m_i^j| \leq \sqrt{2}\sigma. \quad (3.7.60)$$

We then consider the centered (w.r.t. their medians) random variables :

$$\tilde{\xi}_i^j = \xi_i^j - m_i^j,$$

and use the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ together with the upper bound (3.7.60) to deduce that :

$$\begin{aligned} m_n^* &= \mathbb{E} \sup_{k \geq n} \gamma_k^2 \|\xi_k\|^2 = \mathbb{E} \sup_{k \geq n} \gamma_k^2 \sum_{j=1}^d \{\xi_k^j\}^2 \\ &\leq \mathbb{E} \sup_{k \geq n} \gamma_k^2 \left[2 \sum_{j=1}^d \{\xi_k^j - m_j^k\}^2 + 2d\sigma^2 \right] \\ &\leq 2d\sigma^2 \gamma_n^2 + 2\mathbb{E} \sup_{k \geq n} \gamma_k^2 \|\tilde{\epsilon}_k\|^2. \end{aligned}$$

We can use Lemma 3.7.1 and deduce that up to a multiplicative universal constant :

$$m_n^* \lesssim 2d\sigma^2 \gamma_n^2 + 2\sigma^2 \mathbb{E} \sup_{k \geq n} \gamma_k^2 \|Z_k\|^2,$$

where each $(Z_k)_{k \geq n}$ are i.i.d. realizations of Gaussian random variables $\mathcal{N}(0, \sigma^2 I_d)$.

We now aim to apply a chaining argument to control the supremum of the empirical process above. To apply Lemma 3.7.2, we define $\mathcal{T}_n := n; +\infty$ and compute the Laplace transform of the chi-square-like random variables :

$$\log \mathbb{E} e^{\lambda[\gamma_k^2 \|Z_k\|^2 - \gamma_j^2 \|Z_j\|^2]} = \frac{d}{2} \log \left(\frac{1 - 2\lambda\gamma_j^2}{1 - 2\lambda\gamma_k^2} \right)$$

We can check that up to a universal multiplicative constant, we have :

$$\forall \lambda \in \mathbb{R}_+ \quad \forall (a, b) \in \mathbb{R}_+ \times \mathbb{R}_+ : \quad \log \frac{1 - a\lambda}{1 - b\lambda} \lesssim \lambda|a - b| + \frac{|a - b|^2 \lambda^2}{1 - \lambda|a - b|}.$$

We are naturally driven to define the pseudo-metric on \mathcal{T}_n by :

$$\forall (i, j) \in \mathcal{T}_n^2 \quad d(i, j) = |\gamma_i^2 - \gamma_j^2|.$$

It remains to upper bound the covering number of \mathcal{T}_n according to d for any radius $\epsilon > 0$. Indeed, when $2\gamma_n^2 \leq \epsilon$, we have $N(\epsilon, \mathcal{T}_n) = 1$ although when $\epsilon \leq 2\gamma_n^2$, we use the rough bound :

$$N(\epsilon, \mathcal{T}_n) \leq \inf \{j \geq n : 2\gamma_j^2 \leq \epsilon\}.$$

In particular, if $\gamma_j = \gamma j^{-\nu}$, we then obtain

$$N(\epsilon, \mathcal{T}_n) \sim \epsilon^{-1/2\nu}.$$

We apply Lemma 3.7.2 and obtain an upper bound for the right hand side of (3.7.61). The first term is proportionnal to γ_n^2 . The other terms lead to the computation of the two integrals (up to some universal multiplicative constants) :

$$\int_0^{\gamma_n^2} \sqrt{\log(\epsilon^{-1})} d\epsilon \quad \text{and} \quad \int_0^{\gamma_n^2} \log(\epsilon^{-1}) d\epsilon$$

The change of variable $\epsilon = e^{-x}$ and an integration by parts leads to an upper bound whose size is

$\log(\gamma_n^{-2})\gamma_n^2$. \square The next Lemma, borrowed from [30] (see Lemma 13.1, Chapter 13), provides a key estimate for the expectation of the supremum of an empirical process indexed by a pseudo metric space (\mathcal{T}, d) . This estimate involves the covering numbers $N(\delta, \mathcal{T})$ associated with the set \mathcal{T} and the pseudo-metric d .

Lemma 3.7.2. Let \mathcal{T} be a separable metric space and $(X_t)_{t \in \mathcal{T}}$ be a collection of random variables such that for some constants $a, v, c > 0$,

$$\log \mathbb{E} e^{\lambda[X_i - X_j]} \leq a\lambda d(i, j) + \frac{v\lambda^2 d^2(i, j)}{2(1 - c\lambda d(i, j))}$$

for all $(i, j) \in \mathcal{T}^2$ and all $0 < \lambda < \{cd(i, j)\}^{-1}$. Then, for any $i_0 \in \mathcal{T}$:

$$\mathbb{E} \sup_{i \in \mathcal{T}} [X_t - X_{i_0}] \leq 3a\delta + 12\sqrt{v} \int_0^{\delta/2} \sqrt{H(u, \mathcal{T})} du + 12c \int_0^{\delta/2} H(u, \mathcal{T}) du \quad (3.7.61)$$

Chapitre 4

Computation of invariant distributions for stochastic McKean-Vlasov equations

Sommaire

4.1	Introduction	170
4.2	Preliminaries	174
4.3	Hypotheses	178
4.4	Main Result and Scheme of Proof	180
4.4.1	Tightness	181
4.4.2	Identification of the limit	182
4.5	Proofs of the results	184
4.5.1	Proof of Proposition 4.4.14	184
4.5.2	Proof of Theorem 4.4.2	190
4.6	Technical results	194

Abstract The McKean-Vlasov equation is a non-linear process defined by a differential equation where the drift depends on the instantaneous law of the process. In this paper, we investigate such dynamics with a particular focus on the steady regime. More precisely, we build a stochastic algorithm to approximate the invariant measure in the case where this probability is unique.

4.1 Introduction

The McKean-Vlasov equation was first introduced by Anatoly Vlasov for describing time evolution of the distribution function of plasma. It was first studied by Henry McKean in 1966 and interpreted as a probabilistic representation of a class of non-linear parabolic equations (see [89]). To define the process, we fix an interaction potential $W : \mathbb{R}^d \rightarrow \mathbb{R}^+$, a confinement potential $V : \mathbb{R}^d \mapsto \mathbb{R}^+$ and consider

$$\partial_t x_t = \Delta x_t + \operatorname{div}(x_t(\nabla V + \nabla_x W * x_t)) \quad (4.1.1)$$

The probabilistic interpretation of this equation consists in the following SDE :

$$dX_t = -\nabla V(X_t)dt - \nabla_x W * \nu_t(X_t)dt + \sigma dB_t \quad (4.1.2)$$

where $(B_t)_{t \geq 0}$ is a standard brownian motion, with $X_0 \in \mathbb{R}$ and $\nu_t = \mathcal{L}(X_t)$. The symbol $*$ denotes the convolution operator defined by

$$f * \nu(x) = \int_{\mathbb{R}^d} f(x, y) \nu(dy)$$

In particular, if $W = 0$ the process (4.1.1) is the Fokker-Planck equation. More generally, W models the interaction between identical particles and this term makes the equation strongly non-linear in the sense that the coefficients of the equation depend on the instantaneous law of the process.

The existence of solutions for (4.1.2) is not a trivial question and we will not adress this topic in the sequel. Let us point out the work of Gartner [55] and Veretennikov [120] where this question is treated with various formulations depending on the regularity of V and W . An important particularity of (4.1.2) is that the solution $(X_t)_{t \geq 0}$ is not a Markov process since the instantaneous dynamics depends on the whole distribution of X_t . However, the couple $(X_t, \nu_t)_{t \geq 0}$ is markovian on $(\mathbb{R}^d, \mathcal{P}(\mathbb{R}^d))$. Those particularities make the study of the process intricate. A way to facilitate the study is to introduce a system of N particles $(X_t^i)_{1 \leq i \leq N}$ that solves

$$dX_t^i = -\nabla V(X_t^i)dt - \frac{1}{N} \left(\sum_{j=1}^N \nabla_x W(X_t^i, X_t^j) \right) dt + \sigma dB_t^i, \quad i = 1, \dots, d \quad (4.1.3)$$

where $X_0^i = x_0 \in \mathbb{R}$ and $\sigma > 0$. The link between (4.1.2) and (4.1.3) comes from considering N going to infinity. Indeed, the above system aims at modelling the interactions of a system of N particles which approaches (4.1.2) when N goes to infinity. More precisely, when N goes to infinity three effects are observed. First, two particles of the system are less and less correlated. As well

the discrete empirical measure of the system

$$\tilde{\nu}_t^N = \frac{1}{N} \sum_{j=1}^N \delta_{X_t^j}$$

tends to be close to the law of one particle. Finally, one particle behaves more and more as a solution of (4.1.2). This phenomenon is the so-called propagation of chaos (see [117]) and it is at the origin of the approximation technique for McKean-Vlasov equation proposed by [29]. Indeed, their technique is based on the fact that propagation of chaos ensures that the law of X_t can be approached by the empirical measure of the particle system. At this stage, we can also remark that this idea relies on a spatial occupation measure (*i.e.* a measure depending on the position of whole particles at time t) whereas we will use a temporal occupation measure in the sequel.

We aim to propose a numerical algorithm able to approximate the invariant measure of the McKean-Vlasov process. The question of the existence of invariant measure for such a process attracted a lot of attention. Let us recall some of the main results obtained. In a serie of papers Hermann and Tugaut proved existence results in the case where W is convex and V is a double-wells landscape (see [64], [66] and [65]). In the case where V and W are strongly convex, Carillo, McCann and Villani [38] proved the unicity of this invariant measure. Those results led also to results regarding the rate of convergence for a process called Vlasov-Fokker-Planck which satisfises a Fokker Planck equation with an added self-interaction term similar to the one of McKean-Vlasov equations. The work of Cattiaux, Guillin and Malrieu [40] investigate the case of granular media where W is a degenerated convex potential. As the well, the recent work of Monmarché [92] where the convergence to equilibrium of Vlasov-Fokker-Planck equation is studied with a quantitative approach. The propagation of chaos led to study the behavior of the system (4.1.3). For instance, the work of Benachour, Roynette, Talay and Valois [17], [18] investigated the particular system in the case where $V = 0$ showing that the empirical measure does not converge for the L^2 -norm. Concentration properties were also exhibited with the work of Bolley, Guillin and Villani [28] for the Wasserstein distance \mathcal{W}_1 . We also point the recent work of Del Moral and Tugaut [43] which deals with the propagation of chaos in general mean-field models.

Defining the sequence of occupation measures $(\nu_n)_{n \geq 1}$ associated to this process, we aim to prove that $(\nu_n)_{n \geq 1}$ converges almost surely towards the unique invariant measure of the McKean-Vlasov process. This characterization presents a numerical advantage because it only requires the simulation of one trajectory of the process making of it an efficient numerical way of approximation. Such techniques were first investigated in the work of Lamberton and Pagès (see [77]) for diffusion processes where results of convergence and rates of convergence are obtained using mainly martingale methods. Those results were extended by [82] to diffusions with a non-Lipschitz drift and by [104] to Levy processes. However McKean-Vlasov process presents a big problem : it is strongly non-linear. We have to use different tools and this is why we will use the work of [23] and [70] where the longtime behavior of processes close to McKean-Vlasov is studied. They are called self-interacting processes.

Self-interacting processes are defined as follows

$$dY_t = \sigma dB_t + \frac{1}{t} \left(\int_0^t f(Y_t, Y_s) ds \right) dt \quad (4.1.4)$$

where $(B_t)_{t \geq 0}$ is a standard brownian motion, with $X_0 \in \mathbb{R}$ and f is a given function. The proximity between (4.1.2) and (4.1.4) is clear because in the first case we have an interaction induced by the law of $(X_t)_{t \geq 0}$ and in the second case it is induced by the empirical measure of $(X_t)_{t \geq 0}$.

The study of self-interacting diffusions goes back to the work of [42] where the almost sure convergence was proved. The work of [46] led to a conjecture stating that Y_t/t converges a.s for a non-negative and compactly supported function f . It was partially proved by [93]. Our work is directly inspired by the work of [23] and [70] where the study of a class of self-interacting diffusions is done with a focus on the convergence of the occupation measure induced by the process. More precisely, the work of Benaim, Ledoux and Raimond [23] consists in proving that the long time behaviour of the occupation measure $\frac{1}{t} \int_0^t \delta_{X_s} dt$ can be precisely related to the asymptotic pseudotrajectory of an explicit semi-flow on the space of probabilities. Using functional inequalities and pseudotrajectory theory, the characterization is done in the case of a compact Riemannian manifold. To extend this work to \mathbb{R}^d , [70] proposed to modify (4.1.4) by adding a confinement potential V allowing to overcome the non-compactness of \mathbb{R}^d . Indeed, compactness was important in the work of [23] and the function V helped to recreate a kind of compactness. Moreover, this confinement potential makes (4.1.4) close to (4.1.2).

Let us consider $(\tilde{X}_t)_{t \geq 0}$ solution to

$$d\tilde{X}_t = -(\nabla V(\tilde{X}_t) - \nabla_x W * \mu_t(\tilde{X}_t)) dt + \sigma dB_t \quad (4.1.5)$$

where V, W are defined as before and $(\mu_t)_{t \geq 0}$, the occupation measure is defined by

$$\mu_t(dx) = \frac{1}{t} \int_0^t \delta_{X_s}(dx) ds.$$

It can be shown that (see [70]) that the set of invariant distributions of $(\tilde{X}_t)_{t \geq 0}$ corresponds to that of the corresponding McKean-Vlasov SDE (4.1.2). The idea is then to investigate some discretized occupation measures of $(\tilde{X}_t)_{t \geq 0}$ to approximate the invariant distributions of the second one.

We now define an Euler scheme of (4.1.5) with a decreasing step-size. Let $(\gamma_n)_{n \geq 1}$ denotes a decreasing sequence of positive numbers such that

$$\lim_{n \rightarrow \infty} \gamma_n = 0.$$

In particular we consider step-size of the form

$$\gamma_n = \gamma n^{-\beta}$$

where $\gamma > 0$ and $\beta \in (0, 1)$ (the reason why β can not be equal to 1 will appear in Proposition 4.5.2).

Set

$$\Gamma_0 = 0, \quad \Gamma_n = \sum_{k=1}^n \gamma_k \quad \text{for } n \geq 1 \quad (4.1.6)$$

which correspond to the times of discretization and satisfy

$$\lim_{n \rightarrow \infty} \Gamma_n = \infty.$$

We define an Euler scheme $(X_n)_{n \geq 0}$ of the diffusion $(\tilde{X}_t)_{t \geq 0}$ solution of (4.1.5) as follows,

$$X_{n+1} = X_n - \gamma_{n+1} \left(\nabla V(X_n) + \frac{1}{\Gamma_n} \sum_{k=0}^{n-1} \gamma_{k+1} \nabla_x W(X_n, X_k) \right) + \sqrt{\gamma_{n+1}} U_{n+1} \quad (4.1.7)$$

where $X_0 \in \mathbb{R}$ and $(U_n)_{n \geq 1}$ is a white noise such that

$$\mathbb{E}(|U_1|^4) < \infty \quad (4.1.8)$$

We then define a sequence of empirical measures $(\nu_n)_{n \geq 1}$ by

$$\nu_n(dx) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \delta_{X_{k-1}}(dx)$$

In particular, this gives us a recursion formula for $(\nu_n)_{n \geq 0}$:

$$\nu_{n+1} = \nu_n + \frac{\gamma_{n+1}}{\Gamma_n} (\delta_{X_{n+1}} - \nu_n) \quad (4.1.9)$$

This recursive form shows that the formula to compute ν_n is given by a stochastic algorithm with step-size $(\frac{\gamma_{n+1}}{\Gamma_n})_{n \geq 1}$.

Remark Due to the computation of the (discrete) integral between 0 and $n-1$, the complexity of the algorithm is clearly not linear. More precisely, except some particular cases (for instance when W is polynomial), the computation of (X_1, \dots, X_n) requires $\sum_{k=1}^n k \simeq n^2$ operations. To reduce

the complexity, we could consider some less-averaged approximations of $\frac{1}{t} \int_0^t \nabla W(X_t, X_s) ds$ (in the spirit of [100] and [101]). Also, remark that

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \delta_{X_{k-1}} \quad \text{and} \quad \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \nabla_x W(X_n, X_k)$$

are some natural approximations of $\frac{1}{t} \int_0^t \delta_{X_s} ds$ and $\frac{1}{t} \int_0^t \nabla_x W(X_t, X_s) ds$. Nevertheless, algorithms with more general weight sequences could be studied (think for instance to $\frac{1}{n} \sum_{k=0}^{n-1} \delta_{X_k}$).

Our aim in this paper is to study the stochastic algorithm defined by (4.1.9) which provides a way to approximate the invariant measure of the McKean-Vlasov process. It will be proved that the sequence $(\nu_n)_{n \geq 1}$ converges almost surely towards the unique invariant measure of the McKean-Vlasov process. The technique to obtain this result is based on a classical two-step approach. First, we prove the tightness of the occupation measure and then we characterize the limit as a global attractor of a semi-flow.

Before outlining the paper, let us give some notations :

- $\|\cdot\|$ refers to the Euclidean norm on \mathbb{R}^d and $\langle \cdot \rangle$ the underlying scalar product.
- For any $x \in \mathbb{R}^d$, $(x)^{\otimes 2}$ denotes the element of $(\mathbb{R}^d)^2$ defined by :

$$(x)_{i,j}^{\otimes 2} = x_i x_j$$

for all $i, j \in \{1, 2\}$.

- For every function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and any $x, y \in \mathbb{R}^d$, we define

$$D^2 f(x)(y)^{\otimes 2} := \sum_{i,j \in \{1,2\}} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) y_i y_j$$

- For any integer n ,

$$\Delta X_n = X_n - X_{n-1}.$$

The paper is divided as follows. Section 4.2 aims at characterizing the invariant measure of the McKean-Vlasov process as the fix point of a specific map defined on the space of probabilities. It also presents the main features about the equation of interest and some of the tools we will use. Section 4.3 introduces the main assumptions used. Section 4.4 presents the main result of the paper and a guideline of all the proofs. Section 4.5 is devoted to the proof of the main result and as explained before, is divided into two parts. First, we prove tightness and then characterize the limit using an *asymptotic pseudotrajectory* approach.

4.2 Preliminaries

In this section, we characterize the invariant measure of McKean-Vlasov process as a fixed point of the application $\Pi : \nu \mapsto \Pi_\nu$ where Π_ν is the invariant measure of the following SDE :

$$dX_t = -(\nabla V(X_t) + \nabla_x W * \nu(X_t)) dt + \sqrt{2} dB_t, \quad \nu \in \mathcal{P}(\mathbb{R}^d) \quad (4.2.10)$$

This equation can be seen as a "linearized" version of (4.1.2) in the sense that the measure ν does not depend on t . We also introduce operators that will play an important role in our work.

The linearized dynamics

From now on we will denote by $(X_t^\nu)_{t \geq 0}$ the process solution of (4.2.10). Assuming the non-explosion of $(X_t^\nu)_{t \geq 0}$ we call $(P_t^\nu)_{t \geq 0}$ the associated semigroup and the infinitesimal generator L_ν is defined for functions $f \in C^2(\mathbb{R}^d)$ by

$$L_\nu f = D^2 f - \langle \nabla V + \nabla_x W * \nu, \nabla f \rangle, \quad (4.2.11)$$

We also introduce the operator Q_ν defined (when it makes sense) by

$$Q_\nu f := \int_0^\infty (P_t^\nu f - \Pi_\nu f) dt, \quad \text{where } f \in C^\infty(\mathbb{R}^d). \quad (4.2.12)$$

Furthermore, Q_ν is a solution of the Poisson equation associated to L_ν i.e. it satisfies

$$L_\nu \circ Q_\nu = Q_\nu \circ L_\nu = -(Id - \Pi_\nu) =: -K_\nu$$

The existence of the operator Q_ν can be obtained if the semigroup P_t^ν converges at a polynomial rate towards Π_ν i.e. if for some $\rho > 1$ we have

$$|P_t^\nu f - \Pi_\nu f| \leq Ct^{-\rho}$$

Long time stability of $(X_t^\nu)_{t \geq 0}$.

To ensure the existence of an invariant measure we rely on the notion of stability.

DEFINITION 4.2.7. We say that a diffusion $(X_t)_{t \geq 0}$ with infinitesimal generator L is strongly f -stable if one can find a positive function $f \in C^2(\mathbb{R}^d)$ satisfying the following conditions :

- $\lim_{x \rightarrow \infty} f(|x|) = \infty$
- there exist two positive constants a and b such that $Lf \leq b - af$.

Remark The stability may be defined in a slightly different way in the literature. In particular, the mean-reverting condition (the second one) may be less constraining. By *strongly*, we thus emphasize the fact the mean-reverting is assumed to have the intensity of the function f ($Lf \leq b - af$).

In particular, when a diffusion is stable it admits an invariant measure (see [68]). A similar proposition as the following one will be used to derive tightness for our process (see Proposition 4.4.14).

PROPOSITION 4.2.11. *If $(X_t)_{t \geq 0}$ is strongly f -stable then $\sup_{t \geq 0} \mathbb{E}(f(X_t)) < \infty$.*

Proof It suffices to consider $g(t, X_t) = e^{at} f(X_t)$ and to use Ito's formula.

$$g(t, X_t) = g(0, X_0) + \int_0^t e^{as} Lf(X_s) ds + \int_0^t a e^{as} f(X_s) ds + M_t$$

where $(M_t)_{t \geq 0}$ is a local martingale. Introducing a sequence of increasing stopping times $(\tau_n)_{n \geq 1}$ and using the stability condition we get

$$\mathbb{E}(g(t \wedge \tau_n, X_{t \wedge \tau_n})) \leq \mathbb{E}(g(0, X_0)) + b \int_0^{t \wedge \tau_n} e^{as} ds.$$

Applying Fatou's lemma we finally obtain

$$\mathbb{E}(f(X_t)) \leq \mathbb{E}(f(X_0))e^{-at} + \frac{b}{a}, \quad t \geq 0$$

which gives the desired result. □

Applying Proposition 4.2.11, the invariant measure Π_ν associated to (4.2.11) exists under some classical assumptions on V and W . This is summarized by the following proposition.

PROPOSITION 4.2.12. *Assume there exist positive constants a, b, c, d such that for any $x, y \in \mathbb{R}^d$,*

$$\begin{aligned} aV(x) - b &\leq \|\nabla V(x)\|^2 \\ cW(x, y) - d &\leq \|\nabla_x W(x, y)\|^2 \end{aligned}$$

Furthermore assume the Hessian of V and W are bounded on \mathbb{R}^d . Then, the diffusion associated to (4.2.10) is strongly F -stable where

$$F_\nu = V + W * \nu$$

As a consequence, the process $(X_t)_{t \geq 0}$ admits an invariant measure. Denoted by Π_ν , this probability measure admits the following representation :

$$\Pi_\nu(dx) := \frac{e^{-V(x) - W * \nu(x)}}{Z_\nu} \lambda(dx) \quad (4.2.13)$$

where Z_ν is the normalization constant.

The measure Π_ν will appear later when we identify the limit of our process. We also point out that the exhibited function F_ν will be used to derive tightness for our process in Proposition 4.4.14.

The invariant measure of McKean-Vlasov

The study of long time behavior McKean-Vlasov equation (4.1.2) is not a trivial question. On this topic, one can find a lot of references. For instance, the long time behavior of this process has been studied by [38] and recently by [92] with a focus on the rate of convergence towards the invariant measure ν^* . In a serie of papers Hermann and Tugaut proved existence results in the case where W is convex and V is a double-wells landscape (see [64], [66] and [65]). In the case where V and W are strongly convex, Carillo, McCann and Villani [38] proved the unicity of this invariant measure.

Since (4.1.2) presents a particular structure, we should define the notion of invariant measure.

DEFINITION 4.2.8. We call ν^* an invariant measure for McKean-Vlasov equation if for all $t \geq 0$ we have

$$\nu^* P_t^{\nu^*} = \nu^*$$

This measure is characterized as a fixed point of the application $\Pi : \nu \mapsto \Pi_\nu$. In particular, the above definition makes sense since

$$\Pi_{\nu^*} P_t^{\nu^*} = \Pi_{\nu^*}, \quad \forall t \geq 0.$$

The method for proving this result relies on the use of Lyapunov function defined as the sum of the entropy and the potential energy :

$$\mathcal{F}(\nu) = \underbrace{\int_{\mathbb{R}^d} V(x) \nu(x) dx + \int_{\mathbb{R}^d} \nu(x) \log(\nu(x)) dx}_{\text{Entropy}} + \underbrace{\frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x, y) \nu(x) \nu(y) dx dy}_{\text{Potential energy}}$$

Intuitively, this function characterizes the free energy of a system. Laws of Physics state that the free energy should decrease and converge towards its minimum. Differentiating this function, one

can prove it is decreasing and thus converges towards an equilibrium point. In particular those equilibrium are proved to be fixed points of $\Pi : \nu \mapsto \Pi_\nu$. However depending on the convexity of V and W , Π may have several equilibrium points. For instance, we have this result of [64] :

THEOREM 4.2.1. *If V'' is a convex function then, there exist exactly three stationary measures.*

For example, if

$$V(x) = \frac{x^4}{4} - \frac{x^2}{2},$$

where for σ sufficiently small two of the invariant measures are localised around the minima of V and the third one is symmetric.

However, if both V and W are strictly convex, uniqueness is guaranteed. This is why will we work in that framework. The following result of [38] summarizes this property :

THEOREM 4.2.2. *Assume that V, W are convex. Assume moreover that $V \geq 0$ is uniformly convex, in the sense that*

$$D^2V \geq \lambda I.$$

Then, there exists a unique minimizer of \mathcal{F} , which also turns out to be the unique stationary state for equation (4.1.2). In particular, this invariant measure is the unique fixed point of the application Π .

Remark As one can see W is just convex but the need for strong convexity for this function will appear later on in our framework.

Spaces and topology

As mentioned earlier, the function V plays the role of a way of confinement which, in a sense, overcome the lack of compactness of the space. In particular, according to Proposition 2.1, this allows us to work on the compact space of probability $\mathcal{P}_\beta(\mathbb{R}^d; V)$ defined by (4.2.14) below.

We denote by $\mathcal{P}(\mathbb{R}^d)$ the space of probability measures on \mathbb{R}^d , $\mathcal{M}(\mathbb{R}^d)$ the space of signed measures on \mathbb{R}^d . We also consider a subspace of $\mathcal{P}(\mathbb{R}^d)$ defined by

$$\mathcal{P}(\mathbb{R}^d; V) := \left\{ \nu \in \mathcal{P}(\mathbb{R}^d); \int_{\mathbb{R}^d} V(y) |\nu|(dy) < \infty \right\}$$

where $|\nu|$ is the variation of ν : $|\nu| := \nu^+ + \nu^-$ defined using the Hahn-Jordan decomposition of ν . Endowing the space with the following norm defined by

$$\|\nu\|_V := \sup_{f \in C^0(\mathbb{R}^d; V); \|f\|_V \leq 1} \left| \int_{\mathbb{R}^d} f d\nu \right|, \quad f \in \mathcal{P}(\mathbb{R}^d, V)$$

makes it a Banach space. The strong topology of $\mathcal{P}(\mathbb{R}^d, V)$ is the trace of the topology of

$$\mathcal{M}(\mathbb{R}^d, V) := \sup_{f \in C^0(\mathbb{R}^d; V); \|f\|_V \leq 1} \left| \int_{\mathbb{R}^d} f d\nu \right|, \quad f \in \mathcal{M}(\mathbb{R}^d, V)$$

and it makes $\mathcal{P}(\mathbb{R}^d, V)$ a complete metric space. In particular, the weak convergence is defined for a larger class than the usual one.

$$\nu_n \rightarrow \nu, \quad \text{if and only if} \quad \int_{\mathbb{R}^d} f d\nu_n \rightarrow_{n \rightarrow \infty} \int_{\mathbb{R}^d} f d\nu \quad \forall f \in C^0(\mathbb{R}^d, V)$$

For any $\beta > 1$, we introduce the subspace

$$\mathcal{P}_\beta(\mathbb{R}^d; V) := \left\{ \nu \in \mathcal{P}(\mathbb{R}^d; V); \int_{\mathbb{R}^d} V(y) \nu(dy) \leq \beta \right\} \quad (4.2.14)$$

The interest for this subspace comes from the following proposition borrowed from [70].

PROPOSITION 4.2.13. *Let $\beta > 1$. The set $\mathcal{P}_\beta(\mathbb{R}^d; V)$ is a weakly compact subset of $\mathcal{P}(\mathbb{R}^d; V)$.*

We also define the V - norm as follows :

$$\|f\|_V = \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{V(x)}$$

and the space of V -bounded continuous functions as follows

$$C^0(\mathbb{R}^d; V) = \{f \in C^0(\mathbb{R}^d); \|f\|_V < \infty\}$$

Similarly, we define

$$C^\infty(\mathbb{R}^d; V) = C^\infty(\mathbb{R}^d) \cap C^0(\mathbb{R}^d; V).$$

4.3 Hypotheses

In this section, we discuss about the hypothesis we need. Of course, our theoretical results will not involve all these hypotheses simultaneously.

$$\begin{aligned}
(\mathbf{H}_F) \left\{ \begin{array}{l}
\text{(i) } V \in C^2(\mathbb{R}^d) \text{ and } V \geq 0, W \in C^2(\mathbb{R}^{2d}) \text{ and } W \geq 0. \\
\text{(ii) There exists a positive constant } M \text{ such that for any } x, y \in \mathbb{R}^d, \\
\qquad \qquad \qquad \langle \nabla V(x), \nabla_x W(x, y) \rangle \geq -M \\
\text{(iii) There exist positive constants } a, b, c, d \text{ such that for any } x, y \in \mathbb{R}^d, \\
\qquad \qquad \qquad \begin{aligned}
aV(x) - b &\leq \|\nabla V(x)\|^2 \\
cW(x, y) - d &\leq \|\nabla_x W(x, y)\|^2
\end{aligned} \\
\text{(iv) } \sup_{x \in \mathbb{R}^d} \|D^2 V(x)\| < \infty \text{ and } \sup_{x, y \in \mathbb{R}^d} \|D^2 W(x, y)\| < \infty. \\
\text{(v) There exist positive constants } C_1, C_2 \text{ such that for any } x, y \in \mathbb{R}^d \\
\qquad \qquad \qquad \|\nabla V(x)\|^2 \leq C_1 V(x) \quad \|\nabla_x W(x, y)\|^2 \leq C_2 W(x, y). \\
\text{(vi) There exists a positive constant } C_3 \text{ such that for any } x, y \in \mathbb{R}^d \text{ we have} \\
\qquad \qquad \qquad W(x, y) \leq C_3(1 + V(x))(1 + V(y))
\end{array} \right.
\end{aligned}$$

The set of assumptions (i), (ii), (iii) and (iv) is used to derive tightness for $(X_n)_{n \geq 0}$. (iii) ensures the mean reverting effect and (iv) the non-explosion of the algorithm. (ii) looks as a condition on the correlation between the actions of V and W . Roughly, it asks the gradients of V and W are in not too much opposite directions. Assumption (v) states that the functions \sqrt{V} and \sqrt{W} are both Lipschitz continuous. (v) will reveal its importance in the proof of Proposition 4.4.2 and it is a necessary assumption for the approximation of an invariant measure. (vi) is a domination condition which basically states that W is upper-bounded by the function V .

The next set of assumptions concerns Q_ν . More precisely, some smoothness conditions and controls of this operator are needed. The precise conditions are listed below.

$$\begin{aligned}
(\mathbf{H}_Q) \left\{ \begin{array}{l}
\text{For any function } f \in C_c^\infty(\mathbb{R}^d), \\
\text{(i) } x \mapsto Q_\nu f(x) \text{ is of class } C^2(\mathbb{R}^d). \\
\text{(ii) } \sup_{x \in \mathbb{R}^d} |Q_\nu(f(x))| < \infty \\
\text{(iii) } \sup_{x \in \mathbb{R}^d} |\nabla Q_\nu(f(x))| < \infty \\
\text{(iv) } \sup_{x \in \mathbb{R}^d} |D_x^2 Q_\nu f(x)| < \infty \\
\text{(v) For every } \nu, \mu \in \mathcal{P}(\mathbb{R}^d), \\
\qquad \qquad \qquad |Q_\nu f(x) - Q_\mu f(x)| \leq C|f(x)|\|\nu - \mu\|_{TV}
\end{array} \right.
\end{aligned}$$

(H_Q) is a technical assumption on the solution of the Poisson equations $f - \Pi_\nu f = L_\nu g$. If the smoothness in the variable x is a somewhat classical problem when ν is a fixed probability (one can think at least to the unidimensional case), the one in ν seems to be more involved. For this reason, we choose here to admit these conditions and to defer some work on the topic for a future paper. Nevertheless, it is worth noting that due to the formulation of the operator Q_ν , the obtention of sufficient conditions which ensures (H_Q) is strongly related to the long time behavior of the process. One can try to use functional inequalities techniques (see [13] for general results and [23]

for $(H_Q(i)), (H_Q(ii))$ and $(H_Q(iii))$ since we consider compactly supported functions) or PDE techniques (see [74] for example). For $(H_Q(v))$, we refer to section 4.2.1 of [70] for details about the application $\nu \mapsto Q_\nu$. We emphasize that reducing the study of Q to this class of function is not trivial and will be proved by Corollary 4.4.1.

4.4 Main Result and Scheme of Proof

This section presents the main result. For convenience of the reader, we give some scheme of the proof.

Main result

The main result of the paper is the following one.

THEOREM 4.4.1. *Assume that (\mathbf{H}_F) , (\mathbf{H}_Q) and (4.1.8) hold. Suppose that V is strictly convex, W is symmetric and strongly convex then the sequence $(\nu_n)_{n \geq 1}$ converges almost surely towards ν^* , the unique fixed point of the application Π . In other words to the unique invariant distribution of (4.1.2).*

Remarks

One could wonder what would be the result if V and W were no longer convex. A beginning of an answer can be found in [19, Section 6] in Corollary 3.9 of [70] where it is proved that the occupation measure converges towards one of the fixed points of the application Π . Another point left aside is the question of the rate of convergence. In the spirit of [104], we could try to use martingale techniques to establish a CLT for the sequence $(\nu_n)_{n \geq 1}$. However, the problem is here more involved due to the non-linearity of the equation, or more precisely to the infinite dimension of the underlying Markov process. The choice of the step-sizes leaves also interesting questions. Indeed, the step-size of the algorithm and the one of the occupation measure could be different. More precisely, we could also define a sequence of empirical measures $(\tilde{\nu}_n)_{n \geq 1}$ by

$$\tilde{\nu}_n(dx) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{X_k}(dx)$$

where (η_k) is a sequence of positive weights and $H_n = \sum_{k=1}^n \eta_k$. In particular, this would give us a recursion formula for $(\tilde{\nu}_n)_{n \geq 0}$:

$$\tilde{\nu}_{n+1} = \tilde{\nu}_n + \frac{\eta_{n+1}}{H_{n+1}} (\delta_{X_{n+1}} - \tilde{\nu}_n)$$

Some adaptations should be made in our proofs. We refer to [82] for details about this generalization.

Scheme of proof

To obtain this type of result, the scheme of proof is classical. To begin we prove the tightness of the empirical measure using a Lyapunov function. Then we consider a continuous-time process associated to the sequence $(\nu_n)_{n \geq 1}$ and prove it is an asymptotic pseudotrajectory for the semi-flow

of an explicit ODE on a subspace of $\mathcal{M}(\mathbb{R}^d)$. Finally, we derive the convergence of $(\nu_n)_{n \geq 1}$ towards the measure ν^* which is unique if V and W are strongly convex. To guide the reader, the following sections detail the main arguments of the proof.

4.4.1 Tightness

To prove the tightness of $(\nu_n)_{n \geq 1}$, we use a Lyapunov function in the spirit of Proposition 4.2.12. The following result will be proved

PROPOSITION 4.4.14. *Assume $(\mathbf{H_F(i)})$, $(\mathbf{H_F(ii)})$, $(\mathbf{H_F(iii)})$ and $(\mathbf{H_F(iv)})$ and (4.1.8) hold, then*

$$\sup_{n \geq 1} \mathbb{E}(V(X_n)) < \infty. \quad (4.4.12)$$

$$\sup_{n \geq 1} \mathbb{E}(W * \nu_n(X_n)) < \infty. \quad (4.4.13)$$

As a consequence, the occupation measure $(\nu_n)_{n \geq 1}$ is a.s tight.

Remark Hypothesis $H_F(v)$ is not used because of the mean reverting effect guaranteed by $H_F(iv)$. As we will see, $H_F(v)$ is useful for the second part of the proof.

Idea of the proof

The first two assertions come from considering the function $F_{\nu_n}(X_n) = V(X_n) + W * \nu_n(X_n)$. Our objective is to obtain a recursion as follows

$$\mathbb{E}(F_{\nu_{n+1}}(X_{n+1}) | \mathcal{F}_n) \leq F_{\nu_n}(X_n)(1 - c\gamma_{n+1}) + C\gamma_{n+1}. \quad (4.4.14)$$

This recursion will lead to

$$\sup_n \mathbb{E}(F_{\nu_n}(X_n)) < \infty,$$

then conclusion will follow from Proposition 4.5.18 by considering the sequence $(\nu_n(w, \sqrt{F_{\nu_n}}))_{n \geq 1}$ and proving that

$$\sup_n \nu_n(w, \sqrt{F_{\nu_n}}) < \infty, \quad a.s$$

which implies the tightness of the sequence $(\nu_n)_{n \geq 1}$.

We can go further and prove the following result.

PROPOSITION 4.4.15. *Assume that $(\mathbf{H_F})$ hold. Then,*

$$\sup_{n \geq 1} \mathbb{E}(V^2(X_n)) < \infty$$

Remark This result will find its purpose when studying the asymptotic pseudotrajectory related to our process. The proof of this result is adapted from [77] and is postponed to the end of the paper. This result is essential to prove Corollary 4.4.1.

4.4.2 Identification of the limit

To prove that $(\nu_n)_{n \geq 1}$ satisfies the behavior of an ODE, we need to introduce a continuous-time process built as an interpolation of the original procedure. To begin with, we recall the recursion satisfied by $(\nu_n)_{n \geq 1}$.

$$\nu_{n+1} = \nu_n + \frac{\gamma_{n+1}}{\Gamma_{n+1}}(\delta_{X_{n+1}} - \nu_n) \quad (4.4.15)$$

$$= \nu_n + \frac{\gamma_{n+1}}{\Gamma_{n+1}}(\Pi_{\nu_n} - \nu_n) + \frac{\gamma_{n+1}}{\Gamma_{n+1}}(\delta_{X_{n+1}} - \Pi_{\nu_n}) \quad (4.4.16)$$

where Π_ν is defined by (4.2.13) and L_ν by (4.2.11). We set

$$h(X_n) = \Pi_{\nu_n} - \nu_n, \quad \epsilon_n = \delta_{X_{n+1}} - \Pi_{\nu_n}.$$

To set the continuous-time setting, for any $t \geq 0$ we define

$$N(t) = \sup\{k \geq 0, t \geq \tau_k\}$$

where $\tau_n = \sum_{k=1}^n \rho_k$ with

$$\rho_k = \frac{\gamma_k}{\Gamma_k}$$

For any $s \geq 0$,

$$\widehat{\nu}_{s+\tau_n} = \nu_n \left(1 - \frac{s}{\rho_{n+1}}\right) + \frac{s}{\rho_{n+1}} \nu_{n+1}, \quad \bar{\nu}_{s+\tau_n} = \nu_n$$

We also set

$$\bar{\epsilon}_{s+\tau_n} = \epsilon_n, \quad \bar{\gamma}_{s+\tau_n} = \gamma_n.$$

With those notations, we are considering the following equation

$$\widehat{\nu}_t = \nu_0 + \int_0^t h(\bar{\nu}_s) ds + \int_0^t \bar{\epsilon}_s ds.$$

Our aim is to prove the following result.

THEOREM 4.4.2. *The function $(\widehat{\nu}_t)_{t \geq 0}$ is a.s an asymptotic pseudo trajectory for Φ where the flow Φ is solution of the ODE :*

$$\begin{cases} \partial_t \Phi(t, \nu) &= F(\Phi(t, \nu)) \\ \Phi(0, \nu) &= \nu \end{cases}$$

where $F : \nu \mapsto -\nu + \Pi_\nu$.

Remark The reader can find in [70] a proof of the existence and unicity of solutions of (4.4.2).

A continuous function $X : \mathbb{R}^+ \rightarrow \mathcal{P}(\mathbb{R}^d, V)$ where $\mathcal{P}(\mathbb{R}^d, V)$ denotes the space of probability measures on \mathbb{R}^d , is called an asymptotic pseudo-trajectory property for Φ if

$$\lim_{t \rightarrow \infty} \left(\sup_{0 \leq s \leq T} d(X(t+s), \Phi(s, X(t))) \right) = 0$$

where $\Phi : \mathbb{R}^+ \times \mathcal{P}(\mathbb{R}^d, V) \rightarrow \mathcal{P}(\mathbb{R}^d, V)$. Note that one should define the distance d as a metric for the space $\mathcal{P}(\mathbb{R}^d, V)$. As defined in [21], choosing $(f_k)_{k \geq 0}$ a sequence of $C^\infty(\mathbb{R}^d, V)$ functions dense in the space of continuous functions, we consider the following distance

$$d(\mu, \nu) = \sum_{k \geq 0} \frac{1}{2^k} |\mu(f_k) - \nu(f_k)|.$$

To characterize asymptotic pseudotrajectories we rely on the following proposition :

PROPOSITION 4.4.16. *If for any $T > 0$ and any function $f \in C^\infty(\mathbb{R}^d, V)$, we have*

$$\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq T} \left| \int_t^{t+s} \bar{\epsilon}_u f du \right| = 0, \quad a.s \quad (4.4.17)$$

Then, the function $\hat{\nu}_t$ is a.s an asymptotic pseudotrajectory for Φ .

Proof We refer the reader to [70].

□

Actually, we can reduce the class of functions to consider. This is the purpose of the following corollary.

COROLLARY 4.4.1. *Assume $\lim_{|x| \rightarrow \infty} V(x) = \infty$ and that there exists some positive constant C such that for any $x, y \in \mathbb{R}^d$ we have*

$$W(x, y) \leq C(1 + V(x)(1 + V(y))).$$

If $\sup_{n \geq 1} \nu_n(V^2) < \infty$ then Proposition 4.4.16 is still true if (4.4.17) holds only for every function $f \in C_c^\infty(\mathbb{R}^d)$.

Remark For the sake of clarity, the proof of this result is postponed to Section 4.6.

Idea of the proof of Proposition 4.4.16

Following the work of [23] and [70], to prove (4.4.17) we use the solution of the Poisson equation Q_ν associated to L_ν . The link between those two operators is a consequence of Ito's formula as we saw. More precisely, we have for $f \in C_c^\infty(\mathbb{R}^d)$,

$$\epsilon_n f = f(X_{n+1}) - \Pi_{\nu_n} f \quad (4.4.18)$$

$$= L_{\nu_n} \circ Q_{\nu_n} f(X_n) \quad (4.4.19)$$

This leads us to consider the following quantity

$$A_n := \frac{1}{\Gamma_n} (Q_{\nu_n}(f(X_{n+1})) - Q_{\nu_n}(f(X_n))) - \frac{1}{\Gamma_n} R_n - \frac{\sqrt{\gamma_{n+1}}}{\Gamma_n} U_{n+1} \nabla_x Q_{\nu_n}(f(X_n))$$

To handle the terms of the r.h.s we rely on the hypotheses H_Q , Proposition 4.4.14 and Proposition 4.5.18.

To conclude the proof we need some convergence properties of $(\Phi_t)_{t \geq 0}$ towards ν^* . More precisely, we require the critical state ν^* to be a global attractor of Φ . Then using Theorem 4.4.2 we will conclude that $(\nu_n)_{n \geq 1}$ converges almost surely. Before starting, we recall the definition of a global attractor.

DEFINITION 4.4.9. A measure $\nu^* \in \mathcal{P}(\mathbb{R}^d)$ is called a global attractor if

$$\lim_{t \rightarrow \infty} d(\Phi(t, \nu), \nu^*) = 0$$

for any measure $\nu \in \mathcal{P}(\mathbb{R}^d)$.

The following proposition extracted from [70] guarantees the convergence of a solution of (4.4.2).

PROPOSITION 4.4.17. *If W is symmetric and for all $y \in \mathbb{R}^d$ the function $x \mapsto V(x) + W(x, y)$ is strictly convex then, $\lim_{t \rightarrow \infty} d(\Phi(t, \nu), \nu^*) = 0$ where ν^* is the unique fixed point of Π . Thus, ν^* is a global attractor for Φ .*

The result we follow directly from Theorem 4.4.1 since the limit of the flow is the unique global attractor.

4.5 Proofs of the results

4.5.1 Proof of Proposition 4.4.14

We begin by proving (4.4.14).

Our analysis is divided into two parts. Separately, mean reverting effects are exhibited for both functions V and $W * \nu$. In view of $H_F(i)$, we can use a Taylor expansion between X_n and X_{n+1} for V . This yields,

$$V(X_{n+1}) = V(X_n) + \langle \nabla V(X_n), \Delta X_{n+1} \rangle + \frac{1}{2} D^2(V(\xi_n))(\Delta X_{n+1})^{\otimes 2}$$

where $\xi_n \in [X_n, X_{n+1}]$. Using $H_F(iv)$ gives

$$\frac{1}{2} D^2(V(\xi_n))(\Delta X_{n+1})^{\otimes 2} \leq C |\Delta X_{n+1}|^2.$$

This yields,

$$\begin{aligned} V(X_{n+1}) &\leq V(X_n) - \gamma_{n+1} \|\nabla V(X_n)\|^2 - \underbrace{\gamma_{n+1} \langle \nabla V(X_n), \nabla_x W * \nu_n(X_n) \rangle}_{A_n} + \Delta M_{n+1} \\ &\quad + C \gamma_{n+1}^2 \|\nabla V(X_n) + \nabla_x W * \nu_n(X_n)\|^2 + C \gamma_{n+1} |U_{n+1}|^2 + \Delta M_{n+1} \end{aligned}$$

where $(\Delta M_n)_{n \geq 1}$ is a martingale increment. Using $H_F(ii)$ deals with A_n ,

$$A_n \leq M \gamma_{n+1}.$$

Finally,

$$\begin{aligned} V(X_{n+1}) &\leq V(X_n) - c\gamma_{n+1}\|\nabla V(X_n)\|^2 + \gamma_{n+1}(C|U_{n+1}|^2 + M) \\ &\quad + C\gamma_{n+1}^2\|\nabla_x W * \nu_n(X_n)\|^2 + \Delta M_{n+1} \end{aligned}$$

Similarly,

$$\begin{aligned} W * \nu_{n+1}(X_{n+1}) - W * \nu_n(X_n) &= \frac{1}{\Gamma_{n+1}} \sum_{k=1}^{n+1} \gamma_k W(X_{n+1}, X_k) - \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k W(X_n, X_k) \\ &= \frac{\gamma_{n+1}}{\Gamma_{n+1}} W(X_{n+1}, X_{n+1}) + \left(\frac{1}{\Gamma_{n+1}} - \frac{1}{\Gamma_n}\right) \sum_{k=1}^n \gamma_k W(X_n, X_k) \\ &\quad + \frac{1}{\Gamma_{n+1}} \sum_{k=1}^n \gamma_k (W(X_{n+1}, X_k) - W(X_n, X_k)) \\ &\leq \frac{\gamma_{n+1}}{\Gamma_{n+1}} W(X_{n+1}, X_{n+1}) - \frac{\gamma_{n+1}}{\Gamma_{n+1}} W * \nu_n(X_n) \\ &\quad + \frac{1}{\Gamma_{n+1}} \sum_{k=1}^n \gamma_k (W(X_{n+1}, X_k) - W(X_n, X_k)) \end{aligned} \quad (4.5.20)$$

In view of $H_F(i)$, we can use a Taylor expansion between X_n and X_{n+1} for W . This yields,

$$\begin{aligned} W(X_{n+1}, X_k) - W(X_n, X_k) &= -\gamma_{n+1}\langle \nabla_x W(X_n, X_k), \nabla V(X_n) + \nabla_x W * \nu_n(X_n) \rangle \\ &\quad + \frac{1}{2} D_x^2(W(\tilde{\xi}_n))(\Delta X_{n+1})^{\otimes 2} \end{aligned}$$

where $\tilde{\xi}_n \in [X_n, X_{n+1}]$. Using $H_F(iv)$ gives

$$\frac{1}{2} D^2(W(\tilde{\xi}_n))(\Delta X_{n+1})^{\otimes 2} \leq C|\Delta X_{n+1}|^2$$

$$\begin{aligned} W(X_{n+1}, X_k) - W(X_n, X_k) &\leq -\gamma_{n+1}\langle \nabla_x W(X_n, X_k), \nabla_x W * \nu_n(X_n) \rangle \\ &\quad - \underbrace{\gamma_{n+1}\langle \nabla_x W(X_n, X_k), \nabla V(X_n) \rangle}_{B_n} \\ &\quad + C\gamma_{n+1}^2(\|\nabla V(X_n)\|^2 + \|\nabla_x W * \nu_n(X_n)\|^2) + \Delta \tilde{M}_n \\ &\quad + C\gamma_{n+1}|U_{n+1}|^2 \end{aligned}$$

where $(\Delta \tilde{M}_n)_{n \geq 1}$ is a martingale increment. Using $H_F(ii)$ leads to

$$B_n \leq M\gamma_{n+1}.$$

This finally gives

$$\begin{aligned}
W(X_{n+1}, X_k) - W(X_n, X_k) &\leq -\gamma_{n+1} \langle \nabla_x W(X_n, X_k), \nabla_x W * \nu_n(X_n) \rangle \\
&+ C\gamma_{n+1}^2 (\|\nabla V(X_n)\|^2 + \|\nabla_x W * \nu_n(X_n)\|^2) \\
&+ \Delta \tilde{M}_n + C\gamma_{n+1}(1 + |U_{n+1}|^2)
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{1}{\Gamma_{n+1}} &\left(\sum_{k=1}^n \gamma_k W(X_{n+1}, X_k) - W(X_n, X_k) \right) \leq -\frac{\gamma_{n+1}}{\Gamma_{n+1}} \sum_{k=1}^n \gamma_k \langle \nabla_x W(X_n, X_k), \nabla_x W * \nu_n(X_n) \rangle \\
&+ C \frac{\gamma_{n+1}^2}{\Gamma_{n+1}} \sum_{k=1}^n \gamma_k (\|\nabla V(X_n)\|^2 + \|\nabla_x W * \nu_n(X_n)\|^2) + C \frac{\gamma_{n+1}}{\Gamma_{n+1}} \sum_{k=1}^n \gamma_k (1 + |U_{n+1}|^2) \\
&\leq -\gamma_{n+1} \|\nabla_x W * \nu_n(X_n)\|^2 + C\gamma_{n+1}^2 \|\nabla V(X_n)\|^2 + C\gamma_{n+1}^2 \|\nabla_x W * \nu_n(X_n)\|^2 \\
&+ C\gamma_{n+1}(|U_{n+1}|^2 + 1) + \Delta \tilde{M}_n
\end{aligned}$$

Recall that $F_{\nu_n}(X_n) = V(X_n) + W * \nu_n(X_n)$. Combining the previous calculations leads to

$$\begin{aligned}
F_{\nu_{n+1}}(X_{n+1}) - F_{\nu_n}(X_n) &\leq -\gamma_{n+1} \|\nabla_x W * \nu_n(X_n)\|^2 - \gamma_{n+1} \|\nabla V(X_n)\|^2 \\
&+ C\gamma_{n+1}^2 \|\nabla_x W * \nu_n(X_n)\|^2 + C\gamma_{n+1}^2 \|\nabla V(X_n)\|^2 \\
&+ C\gamma_{n+1}(|U_{n+1}|^2 + 1) + \Delta \bar{M}_{n+1}
\end{aligned}$$

where $(\Delta \bar{M}_n)_{n \geq 1}$ is a martingale increment. For n large enough, we have

$$\begin{aligned}
\gamma_{n+1}^2 \|\nabla_x W * \nu_n(X_n)\|^2 - \gamma_{n+1} \|\nabla_x W * \nu_n(X_n)\|^2 &\leq -\tilde{C}\gamma_{n+1} \|\nabla_x W * \nu_n(X_n)\|^2 \\
\gamma_{n+1}^2 \|\nabla V(X_n)\|^2 - \gamma_{n+1} \|\nabla V(X_n)\|^2 &\leq -\tilde{C}\gamma_{n+1} \|\nabla V(X_n)\|^2
\end{aligned}$$

For n large enough, this yields

$$\begin{aligned}
F_{\nu_{n+1}}(X_{n+1}) - F_{\nu_n}(X_n) &\leq -\gamma_{n+1} \|\nabla_x W * \nu_n(X_n)\|^2 - \gamma_{n+1} \|\nabla V(X_n)\|^2 \\
&+ C\gamma_{n+1}(|U_{n+1}|^2 + 1) + \Delta \bar{M}_{n+1}
\end{aligned}$$

Taking the conditional expectation and using (4.1.8) we obtain, for n large enough,

$$\mathbb{E}(F_{\nu_{n+1}}(X_{n+1}) | \mathcal{F}_n) \leq F_{\nu_n}(X_n)(1 - C\gamma_{n+1}) + \tilde{C}\gamma_{n+1}$$

From this inequality we can derive (4.4.12) using the following lemma.

Lemma 4.5.1. Assume that $\gamma_n = \frac{\gamma}{n^\beta}$ where $\beta \in (0, 1]$ and $\gamma > 0$. If for some positive constants a, b we have

$$\mathbb{E}(F_{\nu_{n+1}}(X_{n+1}) | \mathcal{F}_n) \leq F_{\nu_n}(X_n)(1 - a\gamma_{n+1}) + b\gamma_{n+1},$$

then

$$\sup_{n \geq 1} \mathbb{E}(F_{\nu_n}(X_n)) < \infty$$

Proof Indeed, the inequality

$$\mathbb{E}(F_{\nu_{n+1}}(X_{n+1}) | \mathcal{F}_n) \leq F_{\nu_n}(X_n)(1 - a\gamma_{n+1}) + b\gamma_{n+1}$$

leads to

$$\mathbb{E}(F_{\nu_{n+1}}(X_{n+1})) \leq \mathbb{E}(F_{\nu_0}(X_0)) \prod_{k=1}^{n+1} (1 - a\gamma_k) + b \sum_{k=1}^{n+1} \gamma_k \prod_{l=k+1}^{n+1} (1 - a\gamma_l) \quad (4.5.21)$$

Applying Lemma 4.6.1 or Lemma 4.6.2 deals with the second term. For the first one, using the classical inequality

$$1 - x \leq \exp(-x), \quad x \geq 0$$

gives successively,

$$\begin{aligned} \prod_{k=1}^{n+1} (1 - a\gamma_k) &\leq \prod_{k=1}^{n+1} \exp(-a\gamma_k) \leq \exp\left(-a \sum_{k=1}^{n+1} \gamma_k\right) \\ &\leq C \exp(-a\Gamma_{n+1}) \end{aligned}$$

We finally obtain

$$\mathbb{E}(F_{\nu_{n+1}}(X_{n+1})) \leq C \mathbb{E}(F_{\nu_0}(X_0)) \exp(-a\Gamma_{n+1}) + \frac{2b}{a} \quad (4.5.22)$$

Since $\Gamma_n \geq 0$, we obtain the desired result.

□

PROPOSITION 4.5.18. *Assume that for some positive constants a, b we have*

$$\mathbb{E}(F_{\nu_{n+1}}(X_{n+1}) | \mathcal{F}_n) \leq F_{\nu_n}(X_n)(1 - a\gamma_{n+1}) + b\gamma_{n+1},$$

where $F_{\nu_n} := V + W * \nu_n$ is a positive function and (\mathbf{H}_F) holds, then

$$\sup_n \nu_n(w, \sqrt{F_{\nu_n}}) < \infty, \quad \text{a.s.}$$

where

$$\nu_n(w, \sqrt{F_{\nu_n}}) := \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_{k+1} \sqrt{F_{\nu_k}}(X_k(w))$$

Proof of Proposition 4.5.18

The following is adapted from [77]. For some positive constants a, b, α, β , using Jensen's Inequality we have

$$\begin{aligned} \mathbb{E}(\sqrt{F_{\nu_{n+1}}(X_{n+1})} | \mathcal{F}_n) &\leq \sqrt{F_{\nu_n}(X_n)(1 - a\gamma_{n+1}) + b\gamma_{n+1}} \\ &\leq \sqrt{F_{\nu_n}(X_n)(1 - \alpha\gamma_{n+1}) + \beta\gamma_{n+1}} \end{aligned}$$

where we used the following inequality,

$$\sqrt{1+x} \leq 1 + \frac{x}{2}$$

which is true if $x \geq -1$. This yields

$$\sqrt{F_{\nu_n}(X_n)} \leq \frac{\sqrt{F_{\nu_n}(X_n)} - \mathbb{E}(\sqrt{F_{\nu_{n+1}}(X_{n+1})}|\mathcal{F}_n)}{\alpha\gamma_{n+1}} + \frac{\beta}{\alpha} \quad (4.5.23)$$

In particular this inequality implies

$$\bar{\nu}_n(w, \sqrt{F_{\nu_n}(X_n)}) \leq \frac{1}{\Gamma_n} \sum_{k=1}^n \left(\sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_{k+1}}(X_{k+1})}|\mathcal{F}_k) \right) + \frac{\beta}{\alpha}$$

Thus, we now want to prove

$$\sup_{n \geq 1} \frac{1}{\Gamma_n} \sum_{k=1}^n \left(\sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_{k+1}}(X_{k+1})}|\mathcal{F}_k) \right) < \infty.$$

Rearranging the sequence we obtain

$$\begin{aligned} \frac{1}{\Gamma_n} \sum_{k=1}^n \left(\sqrt{F_{\nu_{k-1}}(X_{k-1})} - \mathbb{E}(\sqrt{F_{\nu_k}(X_k)}|\mathcal{F}_{k-1}) \right) &= \frac{1}{\Gamma_n} \sum_{k=1}^{n-1} \left(\sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_k}(X_k)}|\mathcal{F}_{k-1}) \right) \\ &\quad + \frac{\sqrt{F_{\nu_0}(X_0)} - \mathbb{E}(\sqrt{F_{\nu_n}(X_n)}|\mathcal{F}_{n-1})}{\Gamma_n} \end{aligned}$$

The second member tends to 0 and we now deal with the first one.

We study the convergence of the martingale $(M_n)_{n \geq 1}$ defined by

$$M_n = \sum_{k=1}^{n-1} \frac{1}{\Gamma_k} \left(\sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_k}(X_k)}|\mathcal{F}_{k-1}) \right), \quad n \geq 1$$

We now bound the quadratic variation of M_n denoted by $\langle M \rangle_{n \geq 1}$,

$$\begin{aligned} \mathbb{E}(\langle M \rangle_n) &\leq C \sum_{k=1}^{n-1} \frac{1}{\Gamma_k^2} (\mathbb{E}(\sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_k}(X_k)}|\mathcal{F}_{k-1}))^2) \\ &\leq C \sum_{k=1}^{n-1} \frac{1}{\Gamma_k^2} \mathbb{E} \left(\left[\sqrt{F_{\nu_k}(X_k)} - \sqrt{F_{\nu_{k-1}}(X_{k-1})} \right]^2 \right) \end{aligned}$$

We now want to prove that

$$\mathbb{E} \left(\left[\sqrt{F_{\nu_k}(X_k)} - \sqrt{F_{\nu_{k-1}}(X_{k-1})} \right]^2 \right) \leq C\gamma_k \mathbb{E}(F_{\nu_k}(X_k))$$

We begin by using the following decomposition

$$\sqrt{F_{\nu_k}(X_k)} - \sqrt{F_{\nu_{k-1}}(X_{k-1})} = \sqrt{F_{\nu_k}(X_k)} - \sqrt{F_{\nu_k}(X_{k-1})} + \sqrt{F_{\nu_k}(X_{k-1})} - \sqrt{F_{\nu_{k-1}}(X_{k-1})}$$

We directly get that

$$\mathbb{E}(|\sqrt{F_{\nu_k}(X_{k-1})} - \sqrt{F_{\nu_{k-1}}(X_{k-1})}|^2) < \infty$$

in view of (4.4.14). We are now dealing with the remainder term. The function $(x, \nu) \mapsto V(x) + W * \nu(x)$ is of class C^2 in the space variable so we get

$$\sqrt{F_{\nu_k}(X_k)} - \sqrt{F_{\nu_k}(X_{k-1})} = \frac{1}{2} F_{\nu_k}^{-1/2}(\xi_k) \langle \nabla_x F_{\nu_k}(\xi_k), \Delta X_k \rangle$$

where $\xi_k \in (X_{k-1}, X_k)$ and

$$\nabla_x F_{\nu_k}(x) = \nabla V(x) + \nabla_x W(x).$$

Using the following inequality

$$\frac{1}{2}(\sqrt{a} + \sqrt{b}) \leq \sqrt{a+b}$$

with $a = V(x)$ and $b = W * \nu(x)$ leads to

$$\frac{\nabla_x F_{\nu}(x)}{\sqrt{F_{\nu}(x)}} \leq 2 \frac{\nabla V(x) + \nabla_x W * \nu(x)}{\sqrt{V(x)} + \sqrt{W * \nu(x)}} = 2 \frac{\nabla V(x)}{\sqrt{V(x)} + \sqrt{W * \nu(x)}} + \frac{\nabla_x W * \nu(x)}{\sqrt{V(x)} + \sqrt{W * \nu(x)}}.$$

Using $H_F(v)$ we get that for any $x \in \mathbb{R}^d$,

$$\sup_{x \in \mathbb{R}^d} \frac{\nabla_x F_{\nu}(x)}{\sqrt{F_{\nu}(x)}} \leq C.$$

We finally get

$$\sqrt{F_{\nu_k}(X_k)} - \sqrt{F_{\nu_{k-1}}(X_{k-1})} \leq C |\Delta X_k| \leq C \gamma_k (\nabla V(X_k) + \nabla_x W * \nu_k(X_k)) + C \sqrt{\gamma_k} |U_k|$$

Using again $H_F(v)$ we get,

$$\mathbb{E}(\langle M \rangle_n) \leq C \sum_{k=1}^n \frac{\gamma_{k+1}}{\Gamma_k^2} \mathbb{E}(F_{\nu_k}(X_k) + |U_1|)$$

Since

$$\sum_{k=1}^n \frac{\gamma_{k+1}}{\Gamma_k^2} \leq C(1 + \sum_{k=1}^n \frac{1}{k^2}) < \infty$$

and

$$\sup_{n \geq 1} \mathbb{E}(F_{\nu_n}(X_n)) < \infty$$

one can conclude that

$$\mathbb{E}(\langle M \rangle_{\infty}) < \infty$$

so that the martingale $(M_n)_{n \geq 0}$ converges. Setting $a_k = \Gamma_k$ and $b_k = \sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_{k+1}}(X_{k+1})} | \mathcal{F}_k)$ one can apply Kronecker's lemma (see Lemma 4.6.3) and conclude that

$$\frac{1}{\Gamma_n} \sum_{k=1}^n \left(\sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_{k+1}}(X_{k+1})} | \mathcal{F}_n) \right) < \infty, \quad a.s. \quad n \rightarrow \infty.$$

Since

$$\bar{\nu}_n(w, \sqrt{F_{\nu_n}(X_n)}) \leq \frac{1}{\Gamma_n} \sum_{k=1}^n \left(\sqrt{F_{\nu_k}(X_k)} - \mathbb{E}(\sqrt{F_{\nu_{k+1}}(X_{k+1})} | \mathcal{F}_k) \right) + \frac{\beta}{\alpha}$$

the result follows. □

4.5.2 Proof of Theorem 4.4.2

In view of Corollary 4.4.1, one has to prove that

$$\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq T} \left| \int_t^{t+s} \bar{\epsilon}_u f du \right| = 0, \quad a.s$$

which is equivalent in our setting to prove the following

$$\limsup_{n \rightarrow \infty} \max_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} \rho_{i+1} \epsilon_i f \right| = 0. \quad (4.5.24)$$

where $f \in C_c^\infty(\mathbb{R}^d)$. To this end we use the solution of the Poisson equation associated to L_ν as follows.

$$\epsilon_n f = L_{\nu_n} \circ Q_{\nu_n} f(X_n) \quad (4.5.25)$$

In view of $H_Q(i)$ we use a Taylor expansion between X_n and X_{n+1} . This leads to

$$Q_{\nu_n}(f(X_{n+1})) - Q_{\nu_n}(f(X_n)) = \langle \nabla_x Q_{\nu_n} f(X_n), \Delta X_{n+1} \rangle + \frac{1}{2} D_x^2(Q_{\nu_n} f(\xi_n)) (\Delta X_{n+1})^{\otimes 2}$$

where $\xi_n \in [X_n, X_{n+1}]$.

$$\begin{aligned} Q_{\nu_n}(f(X_{n+1})) - Q_{\nu_n}(f(X_n)) &= \langle \nabla_x Q_{\nu_n} f(X_n), \Delta X_{n+1} \rangle + \frac{1}{2} D_x^2(Q_{\nu_n} f(X_n)) (\Delta X_{n+1})^{\otimes 2} \\ &+ \frac{1}{2} D_x^2(Q_{\nu_n} f(\xi_n)) (\Delta X_{n+1})^{\otimes 2} - D_x^2(Q_{\nu_n} f(X_n)) (\Delta X_{n+1})^{\otimes 2} \\ &= \gamma_{n+1} L_{\nu_n} \circ Q_{\nu_n}(f(X_n)) + R_n + \sqrt{\gamma_{n+1}} U_{n+1} \nabla_x Q_{\nu_n}(f(X_n)) \end{aligned}$$

where

$$\begin{aligned} |R_n| &= \left| \frac{1}{2} D_x^2(Q_{\nu_n} f(\xi_n)) (\Delta X_{n+1})^{\otimes 2} - D_x^2(Q_{\nu_n} f(X_n)) (\Delta X_{n+1})^{\otimes 2} \right| \\ &\leq C |\Delta X_{n+1}|^2 \\ &\leq C \gamma_{n+1}^2 (\|\nabla V(X_n)\|^2 + \|\nabla_x W * \nu_n(X_n)\|^2) + C \gamma_{n+1} |U_{n+1}|^2 \end{aligned}$$

where we used $H_Q(iv)$. Now using $H_F(v)$ we obtain

$$|R_n| \leq C \gamma_{n+1}^2 (V(X_n) + W * \nu_n(X_n)) + C \gamma_{n+1} |U_{n+1}|^2 \quad (4.5.26)$$

In view of (4.5.24) and the above computations, we have to handle the following quantity :

$$A_n := \frac{1}{\Gamma_n} (Q_{\nu_n}(f(X_{n+1})) - Q_{\nu_n}(f(X_n))) - \frac{1}{\Gamma_n} R_n - \frac{\sqrt{\gamma_{n+1}}}{\Gamma_n} U_{n+1} \nabla_x Q_{\nu_n}(f(X_n))$$

Using the telescopic expression :

$$\begin{aligned} \frac{1}{\Gamma_n} (Q_{\nu_n}(f(X_{n+1})) - Q_{\nu_n}(f(X_n))) &= \left(\frac{1}{\Gamma_{n+1}} Q_{\nu_{n+1}}(f(X_{n+1})) - \frac{1}{\Gamma_n} Q_{\nu_n}(f(X_n)) \right) \\ &- \left(\frac{1}{\Gamma_{n+1}} Q_{\nu_{n+1}}(f(X_{n+1})) - \frac{1}{\Gamma_n} Q_{\nu_{n+1}}(f(X_{n+1})) \right) \\ &- \left(\frac{1}{\Gamma_n} Q_{\nu_{n+1}}(f(X_{n+1})) - \frac{1}{\Gamma_n} Q_{\nu_n}(f(X_{n+1})) \right) \end{aligned}$$

we can decompose A_n as follows

$$A_n = A_n^1 - A_n^2 - A_n^3 - A_n^4 \quad (4.5.27)$$

where

$$\begin{aligned} A_n^1 &= \frac{1}{\Gamma_{n+1}} Q_{\nu_{n+1}}(f(X_{n+1})) - \frac{1}{\Gamma_n} Q_{\nu_n}(f(X_n)) \\ A_n^2 &= \frac{1}{\Gamma_{n+1}} Q_{\nu_{n+1}}(f(X_{n+1})) - \frac{1}{\Gamma_n} Q_{\nu_{n+1}}(f(X_{n+1})) \\ A_n^3 &= \frac{1}{\Gamma_n} Q_{\nu_{n+1}}(f(X_{n+1})) - \frac{1}{\Gamma_n} Q_{\nu_n}(f(X_{n+1})) \\ A_n^4 &= \frac{1}{\Gamma_n} R_n + \frac{\sqrt{\gamma_{n+1}}}{\Gamma_n} U_{n+1} \nabla_x Q_{\nu_n}(f(X_n)). \end{aligned} \quad (4.5.28)$$

Let us deal with each term separately. To deal with A_n^1 it suffices to notice that it generates a telescopic sum leading to

$$\sum_{i=n}^{k-1} A_i^1 = \sum_{i=n}^{k-1} \left(\frac{1}{\Gamma_{i+1}} Q_{\nu_{i+1}}(f(X_{i+1})) - \frac{1}{\Gamma_i} Q_{\nu_i}(f(X_i)) \right) = \frac{1}{\Gamma_{m+1}} Q_{\nu_{m+1}}(f(X_{m+1})) - \frac{1}{\Gamma_n} Q_{\nu_n}(f(X_n))$$

Using $H_Q(ii)$ we obtain

$$\limsup_{n \rightarrow \infty} \max_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} A_i^1 \right| = 0 \quad (4.5.29)$$

For A_n^2 , we use $H_Q(iii)$ to obtain that

$$|A_n^2| \leq C \left(\frac{1}{\Gamma_n} - \frac{1}{\Gamma_{n+1}} \right) |Q_{\nu_{n+1}}(f(X_{n+1}))| \leq C \frac{\gamma_{n+1}}{\Gamma_n^2}.$$

Since

$$\sum_{n \geq 0} \frac{\gamma_{n+1}}{\Gamma_n^2} < \infty$$

one can conclude that

$$\limsup_{n \rightarrow \infty} \max_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} A_i^2 \right| = 0 \quad (4.5.30)$$

For A_n^3 , using $H_Q(v)$ we obtain

$$|A_n^3| \leq C \frac{1}{\Gamma_n} |f(X_n)| \|\nu_{n+1} - \nu_n\|_{TV} \leq C |f(X_n)| \frac{\gamma_{n+1}}{\Gamma_n^2}$$

In particular, for any $\delta > 0$ we have

$$\mathbb{P} \left(\max_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} A_i^3 \right| \geq \delta \right) \leq \frac{\sum_{n < k \leq N(n, T)} \mathbb{E}(|f(X_k)|) \frac{\gamma_k}{\Gamma_k^2}}{\delta}$$

Recall that the serie $\sum_{n \geq 1} \frac{\gamma_{n+1}}{\Gamma_n^2}$ converges since

$$\sum_{n \geq 1} \frac{\gamma_{n+1}}{\Gamma_n^2} \leq C \int_0^\infty (1+x)^{-2+\beta} dx$$

the r.h.s is finite if and only if $\beta \in (0, 1)$. Using $\sup_{n \geq 1} \mathbb{E}(f(X_n)) < \infty$, we now conclude by Borel-Cantelli's lemma that a.s,

$$\limsup_{n \rightarrow \infty} \max_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} A_i^3 \right| = 0 \quad (4.5.31)$$

Now to take care of the last term. We begin by using (4.5.26) and (4.5.28) to get the following decomposition :

$$A_n^4 \leq C \frac{\gamma_{n+1}^2}{\Gamma_{n+1}} (V(X_n) + W * \nu_n(X_n)) + C \frac{\gamma_{n+1}}{\Gamma_{n+1}} |U_{n+1}|^2 + \frac{\sqrt{\gamma_{n+1}}}{\Gamma_n} U_{n+1} \nabla_x Q_{\nu_n}(f(X_n))$$

First of all, for any $\delta > 0$ we have

$$\mathbb{P} \left(\max_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} \frac{\gamma_{i+1}^2}{\Gamma_{i+1}} (V(X_i) + W * \nu_i(X_i)) \right| \geq \delta \right) \leq \frac{\sum_{n < i \leq N(n, T)} \frac{\gamma_i^2}{\Gamma_i} \mathbb{E}(|V(X_i) + W * \nu_i(X_i)|)}{\delta}$$

Recall that the serie $\sum_{n \geq 1} \frac{\gamma_n^2}{\Gamma_n}$ converges since

$$\sum_{n \geq 1} \frac{\gamma_n^2}{\Gamma_n} \leq C \int_0^\infty (1+x)^{-1-\beta} dx$$

the r.h.s is finite for any $\beta \in (0, 1]$. Using Proposition 4.4.14, we now conclude by Borel-Cantelli's lemma that a.s,

$$\limsup_{n \rightarrow \infty} \max_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} \frac{\gamma_{i+1}^2}{\Gamma_{i+1}} (V(X_i) + W * \nu_i(X_i)) \right| = 0 \quad (4.5.32)$$

To deal with the two other terms, we rely on Proposition 4.2 of [19]. Indeed, since the sequence $(U_n)_{n \geq 0}$ has a fourth order moment (in view of (4.1.8)) and since

$$\sum_{n \geq 0} \frac{\gamma_{n+1}^2}{\Gamma_{n+1}^2} \leq C \sum_{n \geq 0} \frac{1}{(n+1)^2} < \infty,$$

we can apply [25] to conclude that

$$\limsup_{n \rightarrow \infty} \max_{n < k \leq N(n, T)} \sum_{i=n}^{k-1} \frac{\gamma_{i+1}}{\Gamma_{i+1}} |U_{n+1}|^2 = 0$$

Finally, for the last term of A_n^4 we use the following proposition adapted from [19].

Lemma 4.5.2. Assume that

$$\sup_{n \geq 1} \mathbb{E}(\|\nabla_x Q_{\nu_n}(f(X_n))\|^2) < \infty \quad (4.5.33)$$

for any function $f \in C_c^\infty(\mathbb{R}^d)$, then

$$\lim_{n \rightarrow \infty} \sup_{n < k \leq N(n, T)} \left\| \sum_{i=n}^{k-1} \frac{\sqrt{\gamma_{i+1}}}{\Gamma_i} U_{i+1} \nabla_x Q_{\nu_i}(f(X_i)) \right\| = 0$$

Proof of Lemma 4.5.2

Indeed, for any integer $i \in \mathbb{N}$, U_{i+1} is independent from X_i . Thus,

$$\mathbb{E}(U_{i+1} \nabla_x Q_{\nu_i}(f(X_i))) = \mathbb{E}(U_{i+1}) \mathbb{E}(\nabla_x Q_{\nu_i}(f(X_i)))$$

In view of (4.5.33) and (4.1.8) we can apply Proposition 4.2 of [19] with $p = 2$.

□

As we can see we also need to be able to obtain the following inequality :

$$\sup_{n \geq 0} \mathbb{E}(\|\nabla_x Q_{\nu_n} f(X_n)\|^2) < \infty \quad (4.5.34)$$

which is a direct consequence of $H_Q(iii)$ since $f \in C_c^\infty(\mathbb{R}^d)$. As a conclusion, we get

$$\lim_{n \rightarrow \infty} \sup_{n < k \leq N(n, T)} \left| \sum_{i=n}^{k-1} A_i^4 \right| = 0 \quad (4.5.35)$$

Sum up of the results We are finally able to prove the theorem resulting from all those computations.

The proof consists in combining all the previous computations. More precisely, we want to prove

$$\limsup_{n \rightarrow \infty} \max_{n+1 \leq m \leq N(n,T)} \left| \sum_{j=n}^m \rho_j \epsilon_j(f) \right| = 0$$

Recall that

$$\rho_n \epsilon_n(f) = A_n.$$

Thus, combining (4.5.29), (4.5.30), (4.5.31) and (4.5.35) we obtain the desired result.

□

Proof of Theorem 4.4.1

By Theorem 4.4.17, ν^* is a global attractor for Φ . We recall that a global attractor is an attractor whose basin of attraction is the entire space. Then, by Theorem 6.9 of [19], every limit set of any pseudo-trajectory is contained in this attractor. By Theorem 4.4.2, we can conclude that $(\nu_n)_{n \geq 1}$ converges a.s towards ν^* .

4.6 Technical results

Lemma 4.6.1. Set $\gamma_n = \gamma n^{-\beta}$ for $\beta \in (0, 1)$ and $a, \gamma > 0$ then

$$\begin{aligned} \text{(i)} \quad & \sum_{k=1}^n \gamma_{k+1}^2 \prod_{l=k+1}^n (1 - a\gamma_l) \leq \frac{2}{a} \gamma_{n+1} \\ \text{(ii)} \quad & \sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l) \leq \frac{2}{a} \end{aligned}$$

Proof

$$\sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l) \leq \sum_{k=1}^n \gamma_{k+1} \exp \left(-a \sum_{l=k+1}^n \gamma_l \right) \quad (4.6.36)$$

since

$$\sum_{l=k+1}^n \gamma_{k+1} \geq \gamma \sum_{l=k+1}^n \int_k^{k+1} (x+1)^{-\beta} dx \geq \frac{\gamma}{1-\beta} ((n+1)^{1-\beta} - (k+1)^{1-\beta})$$

we have

$$\sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l) \leq \exp \left(-\frac{a\gamma}{1-\beta} (n+1)^{1-\beta} \right) \sum_{k=1}^n \frac{\gamma}{(k+1)^\beta} \exp \left(\frac{a\gamma}{1-\beta} (k+1)^{1-\beta} \right)$$

The function $x \mapsto \frac{\gamma}{x^\beta} \exp \left(\frac{a\gamma}{1-\beta} (x+1)^{1-\beta} \right)$ is non-decreasing on $((\frac{\beta}{a\gamma})^{1-\beta}, \infty)$. Since we can not

ensure this non-decreasing property for all $x \geq 0$, we propose the crude following bound

$$\begin{aligned} \sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l) &\leq 2 \exp\left(-\frac{a\gamma}{1-\beta}(n+1)^{1-\beta}\right) \int_1^n \frac{\gamma}{(x+1)^\beta} \exp\left(\frac{a\gamma}{1-\beta}(x+1)^{1-\beta}\right) dx \\ &\leq \frac{2}{a} \end{aligned}$$

where the last line comes from

$$\int \frac{\gamma}{(x+1)^\beta} \exp\left(\frac{a\gamma}{1-\beta}(x+1)^{1-\beta}\right) dx \leq \frac{1}{a} \exp\left(\frac{a\gamma}{1-\beta}(x+1)^{1-\beta}\right)$$

□

Lemma 4.6.2. Set $a > 0$, $\gamma_n = \frac{\gamma}{n}$ with $\gamma > 0$, we have

$$\sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l) \leq \frac{2}{a}$$

Proof The proof comes directly from the one done for $\beta \in (0, 1)$.

$$\sum_{k=1}^n \gamma_{k+1} \prod_{l=k+1}^n (1 - a\gamma_l) \leq \sum_{k=1}^n \gamma_{k+1} \exp\left(-a \sum_{l=k+1}^n \gamma_l\right) \quad (4.6.37)$$

$$\leq 2 \sum_{k=1}^n \frac{\gamma}{(k+1)} n^{-a\gamma} (k+1)^{a\gamma} \quad (4.6.38)$$

$$\leq \frac{2}{a} \quad (4.6.39)$$

this concludes the proof of the lemma.

□

Lemma 4.6.3. Let $(a_n)_{n \geq 1}$ and $(b_n)_{n \geq 1}$ be two sequences of real numbers. Assume that $(b_n)_{n \geq 1}$ is non-decreasing, positive and satisfies

$$\lim_{n \rightarrow \infty} b_n = \infty$$

Furthermore, if $\sum_n \frac{a_n}{b_n}$ converges in \mathbb{R} then

$$\frac{1}{b_n} \sum_{k=1}^n a_k = 0$$

PROPOSITION 4.6.19. Assume that (\mathbf{H}_F) holds. Then, we have

$$\sup_{n \geq 1} \mathbb{E}(V^2(X_n)) < \infty$$

Proof We start by using a Taylor expansion for the function V^2 between X_{n+1} and X_n .

$$V^2(X_{n+1}) - V^2(X_n) = 2V(X_n)\langle \nabla V(X_n), \Delta X_{n+1} \rangle + \frac{1}{2}D^2(V^2)(\xi_{n+1}).(\Delta X_{n+1})^{\otimes 2}$$

Since

$$D^2(V^2) = 2VD^2V + 2\|\nabla V\|^2$$

In particular, we have

$$D^2(V^2) = 2V(D^2V + \frac{\|\nabla V\|^2}{V})$$

Using $H_F(v)$, we can claim there exists a positive constant C such that

$$D^2(V^2) \leq CV$$

this leads to

$$V^2(X_{n+1}) - V^2(X_n) \leq 2V(X_n)\langle \nabla V(X_n), \Delta X_{n+1} \rangle + CV(\xi_{n+1}).(\Delta X_{n+1})^{\otimes 2} \quad (4.6.40)$$

To deal with $V(\xi_{n+1})$, we use $H_F(iv)$ and obtain

$$V(\xi_{n+1}) = \sqrt{V(\xi_{n+1})^2} \leq (\sqrt{V(X_n)} + L\Delta X_{n+1})^2 \quad (4.6.41)$$

$$= V(X_n) + 2L\sqrt{V(X_n)}\Delta X_{n+1} + L^2|\Delta X_{n+1}|^2 \quad (4.6.42)$$

for some positive constant L . As a consequence, we get

$$\begin{aligned} V^2(X_{n+1}) - V^2(X_n) &\leq 2V(X_n)\langle \nabla V(X_n), \Delta X_{n+1} \rangle + CV(X_n).(\Delta X_{n+1})^{\otimes 2} \\ &\quad + \tilde{C}\sqrt{V(X_n)}.|\Delta X_{n+1}|^3 + c|\Delta X_{n+1}|^4 \end{aligned} \quad (4.6.43)$$

To handle all the terms we must be able to upper bound $|\Delta X_{n+1}|^p$ for $p = 3, 4$. We have

$$\Delta X_{n+1} = -\gamma_{n+1}(\nabla V(X_n) + \nabla W * \nu_n(X_n)) + \sqrt{\gamma_{n+1}}U_{n+1}$$

We will use the following inequality, true for any $p \geq 1$, any $a, b \geq 0$

$$(a + b)^p \leq a^p + p(a + b)^{p-1}b \leq a^p + p2^{p-1}(a^{p-1}b + b^p)$$

which can be proved directly using Newton's formula. For $p = 3$, we have

$$|\Delta X_{n+1}|^3 \leq \gamma_{n+1}^3 |\nabla V(X_n) + \nabla W * \nu_n(X_n)|^3 + C(\gamma_{n+1}^{5/2}|U_{n+1}| \cdot |\nabla V(X_n) + \nabla W * \nu_n(X_n)|^2 + \gamma_{n+1}^{3/2}|U_{n+1}|^3)$$

since $\mathbb{E}(U_{n+1}) = 0$ we are left to deal with $|\nabla V(X_n) + \nabla W * \nu_n(X_n)|^3$.

$$|\nabla V(X_n) + \nabla W * \nu_n(X_n)|^3 \leq |\nabla V(X_n)|^3 + C|\nabla V(X_n)|^2 \cdot |\nabla W * \nu_n(X_n)| + C|\nabla W * \nu_n(X_n)|^3$$

Looking at (4.6.43), we have to upper bound the term

$$\sqrt{V(X_n)}|\nabla V(X_n) + \nabla W * \nu_n(X_n)|^3$$

We have,

$$\begin{aligned}\sqrt{V(X_n)}|\nabla V(X_n) + \nabla W * \nu_n(X_n)|^3 &\leq \sqrt{V(X_n)}|\nabla V(X_n)|^3 \\ &+ C\sqrt{V(X_n)}|\nabla V(X_n)|^2|\nabla W * \nu_n(X_n)| \\ &+ C\sqrt{V(X_n)}|\nabla W * \nu_n(X_n)|^3\end{aligned}$$

Now, we use the inequality

$$ab \leq a^2 + b^2$$

for the second and the third term. This gives successively

$$\begin{aligned}\sqrt{V(X_n)}|\nabla V(X_n)|^2|\nabla W * \nu_n(X_n)| &= \sqrt{V(X_n)}|\nabla W * \nu_n(X_n)||\nabla V(X_n)|^2 \\ &\leq V(X_n)|\nabla W * \nu_n(X_n)|^2 + |\nabla V(X_n)|^4 \\ &\leq V^2(X_n) + |\nabla W * \nu_n(X_n)|^4 + |\nabla V(X_n)|^4\end{aligned}$$

and

$$\sqrt{V(X_n)}|\nabla W * \nu_n(X_n)|^3 \leq \sqrt{V(X_n)}|\nabla W * \nu_n(X_n)||\nabla W * \nu_n(X_n)|^2 \leq C(V^2(X_n) + |\nabla W * \nu_n(X_n)|^4)$$

Now it suffices to use hypothesis $H_F(v)$ to obtain

$$|\nabla W * \nu_n(X_n)|^2 \leq C(W * \nu_n(X_n))$$

Thus,

$$\begin{aligned}\sqrt{V(X_n)}|\Delta X_{n+1}|^3 &\leq C\gamma_{n+1}^3(V^2(X_n) + |W * \nu_n(X_n)|^2) + \gamma_{n+1}^{3/2}|U_{n+1}|^3V(X_n) \\ &+ C\gamma_{n+1}^{5/2}U_{n+1}|\nabla V(X_n) + \nabla W * \nu_n(X_n)|^2\end{aligned}$$

For $p = 4$, the technique is similar

$$\begin{aligned}|\nabla V(X_n) + \nabla W * \nu_n(X_n)|^4 &\leq |\nabla V(X_n)|^4 + C|\nabla V(X_n)|^3|\nabla W * \nu_n(X_n)| + C|\nabla W * \nu_n(X_n)|^4 \\ &\leq |\nabla V(X_n)|^4 + C(|\nabla V(X_n)|^4 + |V(X_n)\nabla W * \nu_n(X_n)|^2) \\ &+ C(|\nabla W * \nu_n(X_n)|^4) \\ &\leq CV^2(X_n) + C|W * \nu_n(X_n)|^2\end{aligned}$$

where we used the inequality

$$ab \leq a^2 + b^2$$

with

$$a = V^2(X_n), \quad b = V(X_n)\nabla W * \nu_n(X_n)$$

The last line comes from hypothesis $H_F(v)$.

As a conclusion we get the following estimate

$$\begin{aligned}
V^2(X_{n+1}) - V^2(X_n) &\leq 2V(X_n)\langle \nabla V(X_n), \Delta X_{n+1} \rangle + CV(X_n) \cdot (\Delta X_{n+1})^{\otimes 2} \\
&+ \tilde{C}\sqrt{V(X_n)}|\Delta X_{n+1}|^3 + c|\Delta X_{n+1}|^4 \\
&\leq -2\gamma_{n+1}V(X_n)|\nabla V(X_n)|^2 \\
&- 2\gamma_{n+1}V(X_n)\langle \nabla V(X_n), \nabla W * \nu_n(X_n) \rangle + \gamma_{n+1}^2(V^2(X_n) \\
&+ W^2 * \nu_n(X_n) + \gamma_{n+1}V(X_n)(|U_{n+1}|^2 + |U_{n+1}|^3))
\end{aligned}$$

Using hypothesis $H_F(iii)$ and $H_F(ii)$ we get

$$-2\gamma_{n+1}V(X_n)|\nabla V(X_n)|^2 - 2\gamma_{n+1}V(X_n)\langle \nabla V(X_n), \nabla W * \nu_n(X_n) \rangle \leq -C\gamma_{n+1}V^2(X_n) + \tilde{C}\gamma_{n+1}V(X_n)$$

In particular, using Holder's inequality it is easy to see that

$$|W * \nu_n(X_n)|^2 \leq C(W^2 * \nu_n(X_n))$$

and now, taking the expectation leads to

$$\begin{aligned}
\mathbb{E}(V^2(X_{n+1})) &\leq \mathbb{E}(V^2(X_n))(1 - C\gamma_{n+1} + C\gamma_{n+1}^2) \\
&- \gamma_{n+1}^2 W^2 * \nu_n(X_n) + C\gamma_{n+1}\mathbb{E}(V(X_n))
\end{aligned}$$

In order to obtain a similar bound for $W^2 * \nu_n(X_n)$, the calculations are similar but we detail the first step. First of all,

$$\begin{aligned}
W^2 * \nu_{n+1}(X_{n+1}) - W^2 * \nu_n(X_n) &= \frac{1}{\Gamma_{n+1}} \sum_{k=1}^{n+1} \gamma_k W^2(X_{n+1}, X_k) - \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k W^2(X_n, X_k) \\
&= \frac{\gamma_{n+1}}{\Gamma_{n+1}} W^2(X_{n+1}, X_{n+1}) \\
&+ \frac{1}{\Gamma_{n+1}} \sum_{k=1}^n \gamma_k (W^2(X_{n+1}, X_k) - W^2(X_n, X_k)) \\
&+ \left(\frac{1}{\Gamma_{n+1}} - \frac{1}{\Gamma_n} \right) \sum_{k=1}^n \gamma_k W^2(X_n, X_k) \\
&\leq \frac{\gamma_{n+1}}{\Gamma_{n+1}} W^2(X_{n+1}, X_{n+1}) \\
&+ \frac{1}{\Gamma_{n+1}} \sum_{k=1}^n \gamma_k (W^2(X_{n+1}, X_k) - W^2(X_n, X_k)) \\
&- \frac{\gamma_{n+1}}{\Gamma_{n+1}} W^2 * \nu_n(X_n)
\end{aligned}$$

Again we have to write a Taylor expansion for the function $W^2(., X_k)$ between X_n and X_{n+1} . The computation is the same as the one done for V^2 so we will not detail it. As a consequence of those

computations we get a recursion as follows :

$$\begin{aligned}\mathbb{E}(V^2(X_{n+1}) + W^2 * \nu_{n+1}(X_{n+1})) &\leq \mathbb{E}(V^2(X_n) + W^2 * \nu_n(X_n))(1 - c\gamma_{n+1} + \tilde{C}\gamma_{n+1}^2) \\ &+ C\gamma_{n+1}\mathbb{E}(V(X_{n+1}) + W * \nu_{n+1}(X_{n+1}))\end{aligned}$$

Thus, for n large enough and using Proposition 4.4.14 we get

$$\mathbb{E}(V^2(X_{n+1}) + W^2 * \nu_{n+1}(X_{n+1})) \leq \mathbb{E}(V^2(X_n) + W^2 * \nu_n(X_n))(1 - \tilde{c}\gamma_{n+1}) + C\gamma_{n+1}$$

In particular, using Lemma 4.5.1 we obtain

$$\sup_{n \geq 1} \mathbb{E}(V^2(X_n)) < \infty$$

□

COROLLARY 4.6.2. *Assume $\mathbf{H_F(vi)}$ holds. If $\sup_{n \geq 1} \nu_n(V^2) < \infty$ then Proposition 4.4.16 is still true if (4.4.17) holds only for every function $f \in C_c^\infty(\mathbb{R}^d)$.*

Proof Let $f \in C^\infty(\mathbb{R}^d, V)$ and consider $\epsilon, M > 0$. We want to prove that

$$\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq T} \left| \int_t^{t+s} \bar{\epsilon}_u f du \right| = 0$$

To this end, we will assume that this equality holds for any function in $C_c^\infty(\mathbb{R}^d)$. Let us notice that f can be approached by a sequence of functions in $C_c^\infty(\mathbb{R}^d)$. Indeed, there exists a sequence $(f_k^M)_{k \geq 1} \subset C_c^\infty(\mathbb{R}^d)$ with $\text{supp}(f_k) \in [-M, M]$ and such that

$$\lim_{k \rightarrow \infty} \|f - f_k^M\|_{\infty, [-M, M]} = 0 \quad (4.6.44)$$

To begin with, we use the following decomposition

$$\int_t^{t+s} \bar{\epsilon}_u f du = \underbrace{\int_t^{t+s} \bar{\epsilon}_u (f - f_k^M) du}_A + \underbrace{\int_t^{t+s} \bar{\epsilon}_u (f_k^M) du}_B \quad (4.6.45)$$

For B , we know that for any $M > 0$ and $k \in \mathbb{N}$

$$\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq T} \left| \int_t^{t+s} \bar{\epsilon}_u (f_k^M) du \right| = 0$$

Let us now consider A . We want to prove that

$$\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq T} \left| \int_t^{t+s} \bar{\epsilon}_u (f - f_k^M) du \right| = 0.$$

which is equivalent to proving that

$$\lim_{n \rightarrow \infty} \sup_{0 \leq s \leq T} \left| \int_{\tau_{N(n,t)}}^{\tau_{N(n,t+s)}} \bar{\epsilon}_u(f - f_k^M) du \right| = 0$$

because of the following inequality

$$\begin{aligned} \left| \int_t^{t+s} \bar{\epsilon}_u(f - f_k^M) du \right| &\leq \left| \int_{\tau_{N(n,t+s)}}^{t+s} \bar{\epsilon}_u(f - f_k^M) du \right| + \left| \int_{\tau_{N(n,t)}}^t \bar{\epsilon}_u(f - f_k^M) du \right| + \left| \int_{\tau_{N(n,t)}}^{\tau_{N(n,t+s)}} \bar{\epsilon}_u(f - f_k^M) du \right| \\ &\leq \left| \int_{\tau_{N(n,t)}}^{\tau_{N(n,t+s)}} \bar{\epsilon}_u(f - f_k^M) du \right| + C|t - \tau_{N(n,t)}| + C|t + s - \tau_{N(n,t+s)}| \\ &\leq \left| \sum_{i=N(n,t)}^{N(n,t+s)+1} \rho_i \bar{\epsilon}_i(f - f_k^M) \right| + 2C\rho_{N(n,t)+1} \end{aligned}$$

We just have now to handle the quantity

$$\left| \sum_{i=N(n,t)}^{N(n,t+s)+1} \rho_i \bar{\epsilon}_i(f - f_k^M) \right| \leq \sum_{i=N(n,t)}^{N(n,t+s)+1} \rho_i |f - f_k^M|(X_{i+1}) + \left| \sum_{i=N(n,t)}^{N(n,t+s)+1} \rho_i \Pi_{\nu_i}(f - f_k^M) \right|.$$

We have the following decomposition

$$\begin{aligned} |f - f_k^M| &\leq |f - f_k^M| 1_{|\cdot| \leq M} + |f - f_k^M| 1_{|\cdot| > M} \\ &\leq \|f - f_k^M\|_{\infty, [-M, M]} + f 1_{|\cdot| > M} \\ &\leq \|f - f_k^M\|_{\infty, [-M, M]} + CV 1_{|\cdot| > M} \end{aligned}$$

Now consider $M > 0$ and $k \in \mathbb{N}$ sufficiently large in order to ensure the following properties :

- $\|f - f_k^M\|_{\infty, [-M, M]} \leq \epsilon.$
- $\sup_{n \geq 1} \nu_n(1_{|\cdot| > M}) \leq \epsilon.$
- $\sup_{n \geq 1} \Pi_{\nu_n}(1_{|\cdot| > M}) \leq \epsilon.$

Thus, we obtain the following upper-bound

$$\sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i |f - f_k^M|(X_{i+1}) \leq \epsilon T + C \sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i V(X_{i+1}) 1_{|X_{i+1}| > M} \quad (4.6.46)$$

Thus, using Cauchy-Schwarz inequality we get

$$\sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i V(X_{i+1}) 1_{|X_{i+1}| > M} \leq \sqrt{\sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i V^2(X_{i+1}) 1_{|X_{i+1}| > M}} \sqrt{\sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i 1_{|X_{i+1}| > M}}$$

In order to get a bound, we need an estimate of $\Gamma_{N(n,s)+1}$. Let us start by noticing that

$$\Gamma_{N(n,t)} - \Gamma_n \simeq T$$

which comes directly from the definition of those quantities. In particular this implies that

$$\sum_{i=n+1}^{N(n,s)} \frac{\gamma_i}{\Gamma_i} \simeq T \quad (4.6.47)$$

Recall that $\gamma n \simeq n^{-\beta}$ with $\beta \in (0, 1)$. As a consequence we get

$$\sum_{i=n+1}^{N(n,s)} \frac{\gamma_i}{\Gamma_i} \simeq \sum_{i=n+1}^{N(n,s)} \frac{1-\beta}{i} \simeq T$$

In view of (4.6.47) we easily get that

$$\sum_{i=n+1}^{N(n,s)} \frac{1-\beta}{i} \simeq (1-\beta) \log\left(\frac{N(n,s)}{n}\right) \simeq T \quad (4.6.48)$$

This gives us

$$\Gamma_{N(n,s)} \simeq \Gamma_{n \exp(\frac{s}{1-\beta})} = \sum_{i=1}^{\lfloor n \exp(\frac{s}{1-\beta}) \rfloor} \simeq C n^{1-\beta} \simeq C \Gamma_n$$

As a consequence we have

$$\sup_{n \geq 1} \frac{\Gamma_{N(n,s)}}{\Gamma_n} < \infty.$$

This leads to the following upper-bound

$$\sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i V(X_{i+1}) 1_{|X_{i+1}| > M} \leq \sqrt{\sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i V^2(X_{i+1})} \sqrt{\sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i 1_{|X_{i+1}| > M}}$$

Since $\sup_{n \geq 1} \frac{\Gamma_{N(n,s)}}{\Gamma_n} < \infty$, we get

$$\begin{aligned} \sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i V(X_{i+1}) 1_{|X_{i+1}| > M} &\leq \frac{1}{\Gamma_n} \sum_{i=n}^{\tilde{N}(n,s)+1} \gamma_i V(X_{i+1}) 1_{|X_{i+1}| > M} \\ &\leq \sqrt{\frac{1}{\Gamma_n} \sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i V^2(X_{i+1})} \sqrt{\frac{1}{\Gamma_n} \sum_{i=n}^{\tilde{N}(n,s)+1} \rho_i 1_{|X_{i+1}| > M}} \\ &\leq \sqrt{\sup_{n \geq 1} \nu_n(V^2)} \sqrt{\sup_{n \geq 1} \nu_n(1_{|\cdot| > M})} \\ &\leq C\epsilon \end{aligned}$$

Similarly, we get

$$\Pi_{\nu_i}(f - f_k^M) \leq C\epsilon + C\Pi_{\nu_i}(V1_{|\cdot|>M})$$

Using Cauchy-Schwarz inequality we get

$$\Pi_{\nu_i}(V1_{|\cdot|>M}) \leq \sqrt{\Pi_{\nu_i}(V^2)}\sqrt{\Pi_{\nu_i}(1_{|\cdot|>M})}$$

Since $W(x, y) \leq C(1 + V(x)(1 + V(y)))$, we get

$$\int_{\mathbb{R}^d} e^{-V(x)-W*\nu_i(x)} dx \geq \int_{\mathbb{R}^d} e^{-CV(x)-\int_{\mathbb{R}^d} V(y)\nu_i(dy)} dx \geq c > 0$$

Thus,

$$\Pi_{\nu_i}(V^2) = \frac{\int_{\mathbb{R}^d} V^2(x)e^{-V(x)-W*\nu_i(x)} dx}{\int_{\mathbb{R}^d} e^{-V(x)-W*\nu_i(x)} dx} dx \quad (4.6.49)$$

$$\leq c^{-1} \int_{\mathbb{R}^d} V^2(x)e^{-V(x)} dx < \infty \quad (4.6.50)$$

where we used the fact that $\lim_{|x| \rightarrow \infty} V(x) = \infty$ and the inequality $x^2 \leq e^{x/2}$ (which is true for x large enough) to conclude that

$$\int_{\mathbb{R}^d} V^2(x)e^{-V(x)} dx < \infty.$$

Recall that the sequence $(\nu_n)_{n \geq 1}$ is tight a.s and the application $\nu \rightarrow \Pi_\nu$ is continuous. As a consequence, the sequence $(\Pi_{\nu_n})_{n \geq 1}$ is a.s tight and we get

$$\Pi_{\nu_i}(1_{|\cdot|>M}) \leq C\epsilon$$

which leads to

$$\Pi_{\nu_i}(V1_{|\cdot|>M}) \leq C\epsilon$$

for M large enough.

Choosing M and k sufficiently large makes ϵ being smaller and smaller, then the conclusion follows. \square

Conclusion et Perspectives

Dans cette thèse, nous avons traité des algorithmes de bandits de Narendra et Shapiro, de la version stochastique de l'algorithme de boule pesante et des processus de McKean-Vlasov avec pour point commun le puissant outil que sont les algorithmes stochastiques. Nous dressons un bilan de nos travaux tout en discutant des possibles pistes de recherche qu'offrent ces thèmes.

Dans l'article *Regret bounds for Narendra and Shapiro bandit algorithms*, nous avons introduit un algorithme de bandit sur-pénalisé et mené une étude du regret à l'aide de technique issues du monde des algorithmes stochastiques. Ce travail a permis de proposer un nouvel algorithme ayant d'excellentes performances vis-à-vis du regret tout étant économe en temps de calcul. Une première extension à ce travail serait de comprendre comment obtenir un résultat similaire dans le cas où le nombre de bras est supérieur ou égal à 3. Les premiers calculs effectués tendent à montrer qu'une telle généralisation passe par une compréhension globale de la dynamique (au sens où il faut travailler sur le comportement de chacun des bras à chaque étape). Une seconde extension serait de généraliser le modèle à des récompenses plus sophistiquées comme des variables aléatoires continues (des variables gaussiennes par exemple) et essayer de donner un sens à l'algorithme avec ce nouveau type de récompense. On pourrait s'intéresser au PMDM qui apparaissait comme limite de l'algorithme renormalisé. En effet, comme nous l'évoquions la mesure limite est peu connue au sens où nous ne connaissons que l'existence et le support de sa densité. Il serait intéressant d'essayer d'affiner notre connaissance de cette mesure en étudiant notamment la régularité de la densité par exemple. Un autre défi pourrait être la recherche de borne inférieure pour la distance de Wasserstein qui nous permettrait d'évaluer la précision de nos bornes.

Dans l'article *Stochastic Heavy Ball*, nous avons étudié une version stochastique de l'algorithme de descente de gradient à mémoire pour lequel nous avons montré la convergence pour plusieurs types de fonctions. Nous avons obtenu la convergence p.s dans le cas d'une fonction coercive régulière puis nous avons précisé ce résultat dans le cas de fonctions fortement convexes en donnant une majoration précise de l'erreur quadratique. Il serait intéressant de généraliser nos méthodes à des algorithmes plus généraux présentant une mémoire. On peut penser aux processus de McKean-Vlasov ou aux modèles de type Cucker-Smale où la mémoire est associée à un terme de dérive et généralise en quelque sorte notre algorithme. Un autre axe de recherche serait la recherche de vitesse de convergence dans le cas non convexe et l'étude de fonctions dont les zéros ne sont pas isolés. En effet, l'un des principaux points qui fait que les théorèmes classiques s'appliquent réside dans le fait que nous supposons les zéros isolés. Les travaux de Vladislav B. Tadic ont ouvert une porte en proposant une étude de la descente de gradient pour des fonctions présentant plusieurs minimums non-isolés. Il serait pertinent de généraliser cette approche à notre cas. Un dernier point sur lequel il faudrait travailler serait le théorème central limite dans le cas de la mémoire polynomiale car pour le moment nous n'avons pas d'idée arrêtée sur la nature de la limite.

Dans la partie *Computation of invariant distributions for stochastic McKean-Vlasov equations*, nous avons proposé un algorithme stochastique permettant d'approcher numériquement la mesure invariante du processus de McKean-Vlasov. Nous avons étendu les méthodes existantes pour les diffusions à des processus non-linéaires afin d'ajuster la procédure. La particularité de notre étude tient dans la façon d'obtenir la convergence des mesures d'occupations. En effet, nous avons exploité la proximité du processus de McKean-Vlasov avec la classe des processus auto-agissants. Nous avons réussi une chose intéressante qui consiste en la réduction de la classe de fonctions à considérer pour l'opérateur Q (nous avons restreint la classe à celle des fonctions C^∞ à support compact). Parmi les extensions possibles, on pense à une étude numérique approfondie en proposant notamment des estimations de la vitesse de convergence de notre procédure et une comparaison avec la mesure limite théorique. Il serait aussi pertinent de se placer dans le cas où la mesure invariante n'est pas unique (en supposant que V et W sont des potentiels dégénérés) afin de comprendre les comportements possibles et essayer de les mettre en avant numériquement. Un travail est en cours pour réduire les hypothèses de notre théorème principal et il semble possible de se passer certaines hypothèses.

Bibliographie

- [1] S. Agrawal and S. Goyal. Analysis of thompson sampling for the multi-armed bandit problem. *JMLR : Workshop and Conference Proceedings vol 23 (2012)*, 23 :39.1–39.26, 2012.
- [2] Z. Allen-Zhu and L. Orecchia. Linear coupling : An ultimate unification of gradient and mirror descent. *Available on arXiv :1407.1537*, 2015.
- [3] J.Y. Audibert and S. Bubeck. Minimax policies for adversarial and stochastic bandits. In *Proceedings of the 22nd Annual Conference on Learning Theory (COLT)*, pages 2635–2686, 2009.
- [4] J.Y. Audibert and S. Bubeck. Regret bounds and minimax policies under partial monitoring. *Journal of Machine Learning Research*, 11 :2635–2686, 2010.
- [5] P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3 :397–422, 2002.
- [6] P. Auer, N. Cesa-Bianchi, Y. Freund, and R.E. Schapire. The non-stochastic multi-armed bandit problem. *SIAM Journal on Computing*, 32 :48–77, 2002.
- [7] P. Auer, R. Ortner, and C. Szepesvari. Improved rates for the stochastic continuum armed bandit problem. *20th Conference on Learning Theory*, 2007.
- [8] B. Awerbuch and H. Kleinberg. Adaptive routing with end to end feedback, distributed learning and geometric approaches. in *Proceedings of the Annual ACM Symposium on Theory of Computing (STOC)*, 2004.
- [9] R. Azais, J.B. Bardet, A. Genadot, N. Krell, and P.A. Zitt. Piecewise deterministic markov process recent results. In *ESAIM Proceedings*.
- [10] F. Bach. Adaptivity of averaged stochastic gradient descent to local strong convexity for logistic regression. *Journal of Machine Learning Research*, 15 :595–627, 2014.
- [11] F. Bach and N. Flammarion. From averaging to acceleration, there is only a step-size. *Proceedings of the International Conference on Learning Theory (COLT)*, 2015.
- [12] F. Bach and M. Moulines. Non-asymptotic analysis of stochastic approximation algorithms for machine learning. *Advances in Neural Information Processing Systems (NIPS)*, 2011.
- [13] D. Bakry, I. Gentil, and M. Ledoux. Analysis and geometry of markov diffusion operators. grundlehren der mathematischen wissenschaften. *Springer*, (2014).
- [14] J.B. Bardet, A. Christen, A. Guillin, F. Malrieu, and P.A. Zitt. Total variation estimates for the TCP process. *Electronic Journal of Probability*, 18 :1–21, 2013.
- [15] J.B. Bardet, A. Christen, A. Guillin, F. Malrieu, and P.A. Zitt. Total variation estimates for the TCP process. *Electron. J. Probab.*, 18 :no. 10, 21, 2013.
- [16] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1) :183–202, 2009.

- [17] S. Benachour, B Roynette, D. Talay, and P. Vallois. Nonlinear self-stabilizing processes. i. existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.* 75(2), (1998).
- [18] S. Benachour, B Roynette, and P. Vallois. Nonlinear self-stabilizing processes. ii. convergence to invariant probability,. *Stochastic Process, Appl.* 75(2), (1998).
- [19] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Seminaire de Probabilités, XXXIII, volume 1709 of Lecture Notes in Math.*
- [20] M. Benaïm. *Dynamics of stochastic approximation algorithms*. Lecture Notes in Mathematics, Séminaire de Probabilités XXXIII. Springer-Verlag, 2006. Characterization and convergence.
- [21] M. Benaïm and M. W. Hirsch. Asymptotic pseudotrajectories and chain recurrent flows, with applications. *J. Dynam. Differential Equation* 8 (1996), 141-176. MR1388167.
- [22] M. Benaïm and M. Hirsh. Asymptotic pseudotrajectories and chain recurrent flows, with applications. *J. Dynam. Differential Equations*, 8 :141–176, 1996.
- [23] M. Benaïm, M. Ledoux, and O. Raimond. Self-interacting diffusions. *Probab. Theory Related Fields* 122 (2002), 1-41. MR1883716.
- [24] M. Benaïm, M. Ledoux, and O. Raimond. Self-interacting diffusions. *Probab. Theory Related Fields*, 122 :1–41, 2002.
- [25] A. Benveniste, M. Métivier, and P Priouret. Adaptive algorithms and stochastic approximations,. volume 22 of *Applications of Mathematics (New York)*. Springer-Verlag, Translated from the French by Stephen S. Wilson., Berlin, 1990. MR1082341.
- [26] B. Bercu and P. Fraysse. A monro for estimation in semiparametric regression models. *Annals of statistics*, 2012.
- [27] P. Billingsley. *Convergence of Probability Measures*. Wiley series in Probability & Statistics, New York, 1995.
- [28] F. Bolley, I. Gentil, and V. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields* 3-4, p.541-593, (2007).
- [29] M. Bossy and D. Talay. A stochastic particle method for some one-dimensional nonlinear pde. *Mathematics and Computer in Simulation*., 38 :43-50, 1995.
- [30] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.
- [31] F. Bouguet, F. Malrieu, F. Panloup, C. Poquet, and J. Reygner. Long time behavior of markov processes and beyond. *ESAIM : Proceedings and Surveys*, pages 193–211, 2015.
- [32] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.
- [33] O. Brandier and M. Duflo. Les algorithmes stochastiques contournent-ils les pièges ? *Annales de l’I.H.P. Probabilités et statistiques*, 32 :395–427, 1996.
- [34] S. Bubeck and N. Cesa-Bianchi. *Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems*, volume 5 of *Foundations and Trends in Machine Learning*. 2012.
- [35] A. Cabot, H. Engler, and S. Gadat. On the long time behavior of second order differential equations with asymptotically small dissipation. *Trans. Amer. Math. Soc.*, 361(11) :5983–6017, 2009.

- [36] A. Cabot, H. Engler, and S. Gadat. Second-order differential equations with asymptotically small dissipation and piecewise flat potentials. In *Proceedings of the Seventh Mississippi State–UAB Conference on Differential Equations and Computational Simulations*, volume 17 of *Electron. J. Differ. Equ. Conf.*, pages 33–38. Southwest Texas State Univ., San Marcos, TX, 2009.
- [37] O. Cappé, A. Garivier, O. Maillard, R. Munos, and G. Stoltz. Kullback-leibler upper confidence bounds for optimal sequential allocation. *Annals of Statistics*, 41 :1516–1 541, 2013.
- [38] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and re- lated equations : entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana* 19 (2003) :no. 3, 971–1018, MR 2053570 (2005a :35126).
- [39] O. Catoni. Rough large deviation estimates for simulated annealing : application to exponential schedules. *Ann. Probab.*, 20(3) :1109â1146, 1992.
- [40] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Probab. Theory Related Fields* 140 (2008) :19–40,no. 1–2, MR 2357669 (2009a :60017).
- [41] B. Cloez. Wasserstein decay of one dimensional jump-diffusions. *preprint*, 2012.
- [42] M. Cranston and Y. Le Jan. Self-attracting diffusions : Two cases studies. *Math. Ann.* 303 (1995), 87–93. MR1348356.
- [43] P. Del Moral and J. Tugaut. Uniform propagation of chaos and creation of chaos for a class of nonlinear diffusions. *Preprint*, 2016+.
- [44] M. Duflo. Random iterative models, adaptive algorithms and stochastic approximations,. *Applications of Mathematics (New York)*. Springer-Verlag, Berlin., 22, 1997.
- [45] F Dufour and Y. Dutuit. Dynamic reliability, a new model. *Proceedings of ESREL 2002 Lambda-Mu 13 Conference*, 2002.
- [46] R. T. Durrett and L. C. G. Rogers. Asymptotic behavior of brownian polymers. *Probab. Theory Related Fields* 92 (1992), 337–349. MR1165516.
- [47] S. Ethier and T. Kurtz. *Markov Processes*. John Willey and Sons, New York, 1986.
- [48] S.N. Ethier and T.G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [49] J. Fontbona, H. Gueïrin, and F. Malrieu. Quantitative estimates for the long time behavior of an ergodic variant of the telegraph process. *Adv. in Appl. Probab.* 44, 2012.
- [50] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval Res. Logist. Quart.*, 3 :95–110, 1956.
- [51] S. Gadat and L. Miclo. Spectral decompositions and l2 operator norms of toy hypocoercive semi groups. *Kinet. Relat. Models* 6, 2013.
- [52] S. Gadat, L. Miclo, and F. Panloup. A stochastic model for speculative bubbles. *Alea : Latin American journal of probability and mathematical statistics*, 12 :491–532, 2015.
- [53] S. Gadat and F. Panloup. Long time behaviour and stationary regime of memory gradient diffusions. *Annales Institut Henri Poincaré (B)*, 50 :564–601, 2014.
- [54] A. Garivier, G. Stoltz, and P. Ménard. Explore first, exploit next, the true shape of regret in bandit problems. *ArXiv :1602.07182*, 2016+.

- [55] J. Gartner. On the mckean-vlasov limit for interacting diffusions. *math. nachr.*, 137, 1988, 197-248.
- [56] E. Ghadimi, H.R. Feyz, and M. Johansson. Global convergence of the heavy-ball method for convex optimization. *arXiv :1412.7457*, 2014.
- [57] S. Ghadimi and G. Lan. Stochastic first- and zeroth-order methods for nonconvex stochastic programming. *SIAM Journal on Optimization*, 23 :2341–2368, 2013.
- [58] S. Ghadimi and G. Lan. Accelerated gradient methods for nonconvex nonlinear and stochastic programming. *Math. Program.*, 156 :59–99, 2016.
- [59] F. Giraud and P. Del Moral. Non-asymptotics analysis of adaptative and annealed feynman-kac particle models. *To appear in Bernoulli Journal*, (2015).
- [60] A. Györfy and G. Ottucska. Adaptive routing using expert advice. *Computer Journal-Oxford*, 49 :180–189, 2006.
- [61] B. Hajek. Cooling schedules for optimal annealing. *Math. Oper. Res.* 13(2), 1988.
- [62] A. Haraux. *Systèmes dynamiques dissipatifs et applications*. R.M.A. Masson, Paris, 1991.
- [63] P. Hartman. *Ordinary Differential Equations*. Classic in Applied Mathematics. Wiley, 1982.
- [64] S. Herrmann and J. Tugaut. Self-stabilizing processes : uniqueness problem for stationary measures and convergence rate in the small noise limit. preprint available on Herrmann’s web page.
- [65] S. Herrmann and J. Tugaut. Stationary measures for self-stabilizing diffusions : asymptotic analysis in the small noise limit. *Elect. Journ. Probab.*, 15 (2010) :2087–2116.
- [66] S. Herrmann and J. Tugaut. Non-uniqueness of stationary measures for self-stabilizing processes. *Stochastic Process. Appl.* 120 (2010) :no. 7, 1215–1246., MR 2639745.
- [67] C. Hu, W. Pan, and J.T Kwok. Accelerated gradient methods for stochastic optimization and online learning. *In Advances in Neural Information Processing Systems*, 2009.
- [68] R. Z. Khas’minskii. Stochastic stability of differential equations, translated from the russian by d. louvish., volume 7 de monographs and textbooks on mechanics of solids and fluids : Mechanics and analysis. *Sijthoff and Noordhoff, Alphen aan den Rijn*, 1980.
- [69] J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, 23 :462–466, 1952.
- [70] A. Kurtzmann. The ode method for some self-interacting diffusions on \mathbb{R}^d . *Annales de l’institut Henri Poincaré* :Vol .46 (3), 618–643, 2010.
- [71] H. Kushner and G. Yin. *Stochastic approximation and recursive algorithms and applications, Second edition*. Applications of Mathematics, 35. Stochastic Modelling and Applied Probability. Springer-Verlag, New-York, 2003.
- [72] H.J. Kushner and D.S. Clark. *Stochastic approximation for constrained and unconstrained systems*. Springer-Verlag, Berlin, 1978.
- [73] H.J. Kushner and G. Yin. Stochastic approximation and recursive algorithms and applications. *Springer-Verlag, second edition*,, 2003.
- [74] O. Ladyzhenskaya and N. Ural’steva. Linear and quasilinear elliptic equation. *Translated from Russian by Scripta Technica, Inc. Translation Editor : Leon Ehrenpreis. Academic Press, New-York London*, 1968.
- [75] T.L. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6 :4–22, 1985.

- [76] D. Lamberton and G. Pagès. How fast is the bandit ? *Stochastic Analysis and Applications*, 26 :603–623, 2008.
- [77] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8 :367–405.
- [78] D. Lamberton, G. Pagès, and P. Tarrès. When can the two-armed bandit algorithm be trusted ? *Annals of Applied Probability*, 14 :1424–1454, 2004.
- [79] D. Lamberton and Gilles Pagès. A penalized bandit algorithm. *Electronic Journal of Probability*, 13 :341–373, 2008.
- [80] G. Lan. An optimal method for stochastic composite optimization. *Math. Program*, 133(1-2, Ser.A) :365–397, 2012.
- [81] J. Lee, M. Simchowitz, M. Jordan, and B. Recht. Gradient descent converges to minimizers. *Preprint*, 2016.
- [82] V. Lemaire. *Estimation numérique de la mesure invariante d’un processus de diffusion. Ph.D. thesis*. PhD thesis, Univ. Marne-La Vallée, (2005).
- [83] V. Lemaire. An adaptive scheme for the approximation of dissipative systems. *Stochastic Processes and their Applications*, 117(10) :1491–1518, 2007.
- [84] M. Lepilliez. Simulation numérique des ballotements d’ergols dans les réservoirs de satellites en microgravité et à faible nombre de bond. *Thèse de doctorat à l’IMFT*, 2015.
- [85] K. Liu and Q. Zhao. A restless bandit formulation of opportunistic access, indexability and index policy. In *Proc. of the 5th IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON) Workshops*, 2008.
- [86] L. Ljung. Strong convergence of a stochastic approximation algorithm. *Annals of statistics*, 1978.
- [87] O.-A. Maillard, R. Munos, and G. Stoltz. A finite-time analysis of multiarmed bandits problems with kullback-leibler divergences. in *Proceedings of the Annual Conference on Learning Theory (COLT)*, 2011.
- [88] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations : locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2) :185–232, 2002.
- [89] H. P. McKean. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967) Air Force Office Sci. Res.*, Arlington, Va , MR 0233437 (38 1759) :pp. 41–57., 1967.
- [90] H. McMahan and A. Blum. Online geometric optimization in the bandit setting against an adaptive adversary. *COLT proceedings*, 2004.
- [91] Sean P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3) :518–548, 1993.
- [92] P. Monmarché. Ergodicity for mean field hamiltonian particles associated to a vlasov-fokker-planck equation. <http://arxiv.org/abs/1603.03179>, 2016.
- [93] T. S. Mountford and P. P. Tarrès. An asymptotic result for brownian polymers. *Ann. Inst. H. Poincaré, Probab. Statist.* 44, (2008) 29–46.
- [94] K. S. Narendra and I. J. Shapiro. Use of stochastic automata for parameter self-optimization with multi-modal performance criteria. *IEEE Trans. Syst. Sci. Cybern.*, 5 :352–360, 1969.

- [95] A. Nemirovski and D. Yudin. Problem complexity and method efficiency in optimization. *Wiley-Interscience Series in Discrete Mathematics.*, John Wiley, XV, 1983.
- [96] N. Nesterov. A method of solving a convex programming problem with convergence rate $o(1/k^2)$. *Soviet Mathematics Doklady*, 27(2) :372–376, 1983.
- [97] Y. Nesterov. *Introductory lectures on convex optimization*, volume 87 of *Applied Optimization*. Kluwer Academic Publishers, Boston, MA, 2004. A basic course.
- [98] M.F. Norman. On linear models with absorbing barriers. *J. Math. Psych.*, 5 :225–241, 1968.
- [99] H. G. Othmer, S. R. Dunbar, and W. Alt. Models of dispersal in biological systems. *J. Math. Biol.* 26, 1988.
- [100] G. Pages and F. Panloup. Approximation of the distribution of a stationary markov process with application to option pricing. *Bernoulli*, 15(1) : 146-177, 2009.
- [101] G. Pages and F. Panloup. Ergodic approximation of the distribution of a stationary diffusion : rate of convergence. *Annals of Applied Probability*, 22(3) : 1059-1100, 2012.
- [102] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Adv. in Appl. Probab.*, 2010.
- [103] S. Pandey, D. Agarwal, D. Chakrabarti, and V. Josifovski. Bandits for taxonomies : A model based approach. In *Proceedings of the Seventh SIAM International Conference on Data Mining*, 2007.
- [104] F. Panloup. Approximation du régime stationnaire d’une eds avec sauts. *Ph.D. thesis, Univ. Paris VI.*, (2006).
- [105] R. Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 1-79, 2007.
- [106] R. Pemantle. Non-convergence to unstable points in urn models and stochastic approximations. *Annals of Probability*, 18 :698–712, 1990.
- [107] V. Perchet and P. Rigollet. The multiarmed bandit problem with covariates. *Annals of statistics*, 2013.
- [108] H. Poincaré. Mémoire sur les courbes définies par une équation différentielle (iv). *Journal de Mathématiques Pures et Appliquées*, 4 :151–217, 1886.
- [109] B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4 :1–17, 1964.
- [110] B.T Polyak and A. Juditsky. Acceleration of stochastic approximation by averaging. *SIAM Journal on Control and Optimization*, 30 :838–855, 1992.
- [111] H. Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematics Society*, 58 :527–535, 1952.
- [112] H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Statist.*, 22 :400–407, 1951.
- [113] D. Ruppert. Efficient estimations from a slowly convergent robbins-monro process. *Technical Report, 781, Cornell University Operations Research and Industrial Engineering*, 1988.
- [114] S. S. Boyd, E.J Candes, and W. Su. A differential equation for modeling nesterov’s accelerated gradient method : theory and insights. *Journal of Machine Learning Research*, to appear, 2016.
- [115] K.R. Stromberg. *Probability for Analysts*. Chapman & Hall, CRC, New York, 1994.

- [116] D.W. Stroock and S. R. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
- [117] A.S. Sznitman. Topics in propagation of chaos. *In École d'Été de Probabilités de Saint-Flour XIX 1989, volume 1464 of Lecture Notes in Math., pages 165–251. Springer, Berlin, 1991.*
- [118] D. Talay. Second order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stoch. Stoch. Rep.*, 29(1) : 13-36, 1990.
- [119] Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the royal Statistical Society, Series B*, 58(1), 267-288, 1996.
- [120] A. Y. Veretennikov. On ergodic measures for mckean-vlasov stochastic equations, monte carlo and quasi-monte carlo methods 2004. *Springer, Berlin, 2006, pp. 471-486. MR 2208726 (2007a :60039).*
- [121] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950), 2009.
- [122] T. Yang, Q. Lin, and Z. Li. Unified convergence analysis of stochastic momentum methods for convex and non-convex optimization. *Preprint*, 2016.

the most profitable without spending too much times on the wrong ones. Our goal is to understand the weakness of this algorithm in order to propose an optimal procedure for a quantity measuring the performance of a bandit algorithm, the regret. In our results, we will propose an algorithm called NS over-penalized which allows to obtain a minimax regret bound. A second work will be to understand the convergence in law of this process. The particularity of the algorithm is that it converges in law toward a non-diffusive process which makes the study more intricate than the standard case. We will use coupling techniques to study this process and propose rates of convergence.

The second work of this thesis falls in the scope of optimization of a function using a stochastic algorithm. We will study a stochastic version of the so-called heavy ball method with friction. The particularity of the algorithm is that its dynamics is based on the all past of the trajectory. The procedure relies on a memory term which dictates the behavior of the procedure by the form it takes. In our framework, two types of memory will be investigated : polynomial and exponential. We will start with general convergence results in the non-convex case. In the case of strongly convex functions, we will provide upper-bounds for the rate of convergence. Finally, a convergence in law result is given in the case of exponential memory.

The third part is about the McKean-Vlasov equations which were first introduced by Anatoly Vlasov and first studied by Henry McKean in order to model the distribution function of plasma. Our objective is to propose a stochastic algorithm to approach the invariant distribution of the McKean Vlasov equation. Methods in the case of diffusion processes (and some more general processes) are known but the particularity of McKean Vlasov process is that it is strongly non-linear. Thus, we will have to develop an alternative approach. We will introduce the notion of asymptotic pseudotrajectory in order to get an efficient procedure.

Dans cette thèse, nous étudions des thématiques autour des algorithmes stochastiques et c'est pour cette raison que nous débuterons ce manuscrit par des éléments généraux sur ces algorithmes en donnant des résultats historiques pour poser les bases de nos travaux. Ensuite, nous étudierons un algorithme de bandit issu des travaux de Narendra et Shapiro dont l'objectif est de déterminer parmi un choix de plusieurs sources laquelle profite le plus à l'utilisateur en évitant toutefois de passer trop de temps à tester celles qui sont moins performantes. Notre but est dans un premier temps de comprendre les faiblesses structurelles de cet algorithme pour ensuite proposer une procédure optimale pour une quantité qui mesure les performances d'un algorithme de bandit, le regret. Dans nos résultats, nous proposerons un algorithme appelé NS sur-pénalisé qui permet d'obtenir une borne de regret optimale au sens minimax au travers d'une étude fine de l'algorithme stochastique sous-jacent à cette procédure. Un second travail sera de donner des vitesses de convergence pour le processus apparaissant dans l'étude de la convergence en loi de l'algorithme NS sur-pénalisé. La particularité de l'algorithme est qu'il ne converge pas en loi vers une diffusion comme la plupart des algorithmes stochastiques mais vers un processus à sauts non-diffusif ce qui rend l'étude de la convergence à l'équilibre plus technique. Nous emploierons une technique de couplage afin d'étudier cette convergence.

Le second travail de cette thèse s'inscrit dans le cadre de l'optimisation d'une fonction au moyen d'un algorithme stochastique. Nous étudierons une version stochastique de l'algorithme déterministe de boule pesante avec amortissement. La particularité de cet algorithme est d'être articulé autour d'une dynamique qui utilise une moyennisation sur tout le passé de sa trajectoire. La procédure fait appelle à une fonction dite de *mémoire* qui, selon les formes qu'elle prend, offre des comportements intéressants. Dans notre étude, nous verrons que deux types de mémoire sont pertinents : les mémoires exponentielles et polynomiales. Nous établirons pour commencer des résultats de convergence dans le cas général où la fonction à minimiser est non-convexe. Dans le cas de fonctions fortement convexes, nous obtenons des vitesses de convergence optimales en un sens que nous définirons. Enfin, l'étude se termine par un résultat de convergence en loi du processus après une bonne renormalisation.

La troisième partie s'articule autour des algorithmes de McKean-Vlasov qui furent introduit par Anatoly Vlasov et étudié, pour la première fois, par Henry McKean dans l'optique de la modélisation de la loi de distribution du plasma. Notre objectif est de proposer un algorithme stochastique capable d'approcher la mesure invariante du processus. Les méthodes pour approcher une mesure invariante sont connues dans le cas des diffusions et de certains autre processus mais ici la particularité du processus de McKean-Vlasov est de ne pas être une diffusion linéaire. En effet, le processus a de la mémoire comme les processus de boule pesante. De ce fait, il nous faudra développer une méthode alternative pour contourner ce problème. Nous aurons besoin d'introduire la notion de pseudo-trajectoires afin de proposer une procédure efficace.

In this thesis, we are studying several stochastic algorithms with different purposes and this is why we will start this manuscript by giving historical results to define the framework of our work. Then, we will study a bandit algorithm due to the work of Narendra and Shapiro whose objective was to determine among a choice of several sources which one is